How can sequences with frameshifts and STOP-CODON be 'productive' in IMGT/(High)V-QUEST?

Example 1.

File 5, AA sequence, filter for productive:

QVQLQQPGAELVKPGASVMLSCKASGYTFTSYWMHWVKQRPGQAQGHTDCRQTLQHSLHAAQQPDI*GLCGLLLCKR*LPLLL*LLGPRHHSHSLL

corresponding File 3, nt sequence:

caggtccaactgcagcagcctggggctgagcttgtgaagcctggggcttcagtgatgctgtcctgcaaggcttctggctacaccttcaccagctactggatgcactgggtgaagcagaggcctggacaagcgcaaggccacactgactgtagacaaaccctccagcacagcctacatgcagctcagcagcctgacatctgaggactctgcggtctattattgtgcaagaggtaactacccctactactttgactactggggccaaggcaccactctcacagtctcctcag

Analysed sequence length: 290.
Sequence analysis category: 2 (indel search & correction).
Sequence compared with the *Mus musculus* (house mouse) IG set from the IMGT reference directory

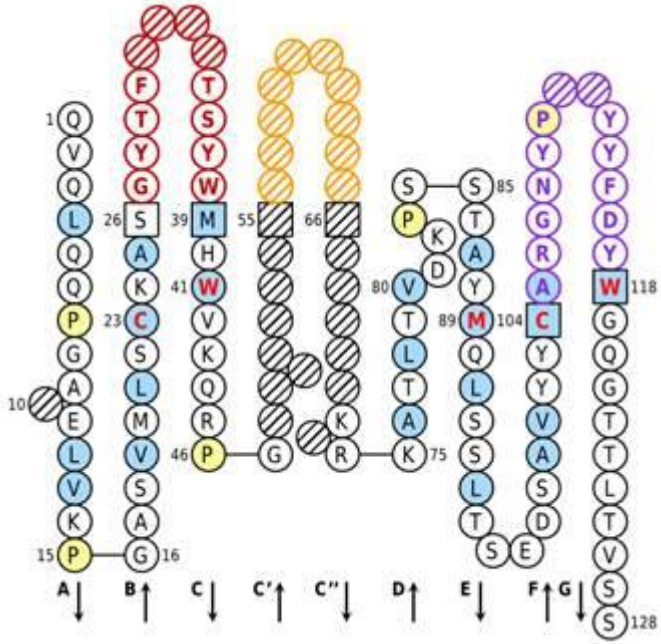
```
>seq1_TW
caggtccaactgcagcagcctggggctgagcttgtgaagcctggggcttcagtgatgctg
tcctgcaaggcttctggctacaccttcaccagctactggatgcactgggtgaagcagagg
cctggacaagcgcaaggccacactgactgtagacaaaccctccagcacagcctacatgca
gctcagcagcctgacatctgaggactctgcggtctattattgtgcaagaggtaactaccc
ctactactttgactactggggccaaggcaccactctcacagtctcctcag
```

Result summary:	**Nucleotide deletions have been detected** (shown by dots in the alignments):			
	localization	nb of deleted nt	causing frameshift	from V-REGION codon	from nt position in user submitt
	FR2-IMGT	65	yes	48	128

IMGT/V-QUEST results after filling the deletion(s) gap(s): Potentially productive IGH rearranged sequence (no stop codon and in-frame junction) (Check also your sequence with BLAST against IMGT/GENE-DB reference sequences to eventually identify out-of-frame pseudogenes)			
V-GENE and allele	Musmus IGHV1-72*01 F	score = 1092	identity = 99.10% (221/223 nt) [98.65% (220/223 nt)]
J-GENE and allele	Musmus IGHJ2*01 F	score = 240	identity = 100.00% (48/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Musmus IGHD2-1*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.9.32.11]	[8.X.11]	CARGNYPYYFDYW

7. V-REGION translation

```
                                 <------------------------------------------- FR1 - IMGT
                                 1               5                   10                  15
                                  Q   V   Q   L   Q   Q   P   G   A       E   L   V   K   P
seq1_TW                          cag gtc caa ctg cag cag cct ggg gct ... gag ctt gtg aag cct

AC163348 Musmus IGHV1-72*01 F    --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---

                                 -----------------------------------------> ______________
                                                 20                  25                  30
                                  G   A   S   V   M   L   S   C   K   A   S   G   Y   T   F
seq1_TW                          ggg gct tca gtg atg ctg tcc tgc aag gct tct ggc tac acc ttc
                                                  K
AC163348 Musmus IGHV1-72*01 F    --- --- --- --- -a- --- --- --- --- --- --- --- --- --- ---

                                 __  CDR1 - IMGT  _____________ <------------------------
                                                 35                  40                  45
                                                  T   S   Y   W   M   H   W   V   K   Q   R
seq1_TW                          ... ... ... ... acc agc tac tgg atg cac tgg gtg aag cag agg

AC163348 Musmus IGHV1-72*01 F    ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---

                                  FR2 - IMGT  ------------------------> _____________  CDR2
                                                 50                  55                  60
                                  P   G
seq1_TW                          cct gga c.. ... ... ... ... ... ... ... ... ... ... ... ...
                                          R   G   L   E   W   I   G   R   I   D   P   N
AC163348 Musmus IGHV1-72*01 F    --- --- -ga ggc ctt gag tgg att gga agg att gat cct aat ...

                                  - IMGT  ________ <---------------------------------------
                                                 65                  70                  75
                                                                              K       R   K
seq1_TW                          ... ... ... ... ... ... ... ... ... ... ... aag ... cgc aag
                                      S   G   G   T   K   Y   N   E   K   F           S
AC163348 Musmus IGHV1-72*01 F    ... agt ggt ggt act aag tac aat gag aag ttc --- ... a-- ---

                                 ------------------------------  FR3 - IMGT  --------------
                                                 80                  85                  90
                                  A   T   L   T   V   D   K   P   S   S   T   A   Y   M   Q
seq1_TW                          gcc aca ctg act gta gac aaa ccc tcc agc aca gcc tac atg cag

AC163348 Musmus IGHV1-72*01 F    --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                 ---------------------------------------------------> ____
                                                 95                  100             104
                                  L   S   S   L   T   S   E   D   S   A   V   Y   Y   C   A
seq1_TW                          ctc agc agc ctg aca tct gag gac tct gcg gtc tat tat tgt gca

AC163348 Musmus IGHV1-72*01 F    --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                 __________  CDR3 -  IMGT  __________

                                  R   G   N   Y   P   Y   Y   F   D   Y   W   G   Q   G   T
seq1_TW                          aga ggt aac tac ccc tac tac ttt gac tac tgg ggc caa ggc acc

AC163348 Musmus IGHV1-72*01 F    ---
```



Example 2.

File 5, AA sequence, filter for productive:

QVQLQQPGAELVKPGASVKL*SRGLDEALSGLEGLIPIVVVLSTMRSSRARPH*L*TNPPAQPTCSSAA*HLRTLRSIIVQDTVTTATLTTGAKAPLSQSPQ

corresponding File 3, nt sequence:

caggtccaactgcagcagcctggggctgagcttgtgaagcctggggcttcagtgaagctgtgaagcagaggcctggacgaggccttgagtggattggaaggattgatcccaatagtggtggtgctaagtacaat
gagaagttcaagagcaaggccgcactgactgtagacaaaccctccagcacagcctacatgcagctcagcagcctgacatctgaggactctgcggtctattattgtgcaagatacggtaactacggctactttgac
tactggggccaaggcaccactctcacagtctcctcag

Analysed sequence length: 306.
Sequence analysis category: 2 (indel search & correction).
Sequence compared with the *Mus musculus* (house mouse) IG set from the IMGT reference directory

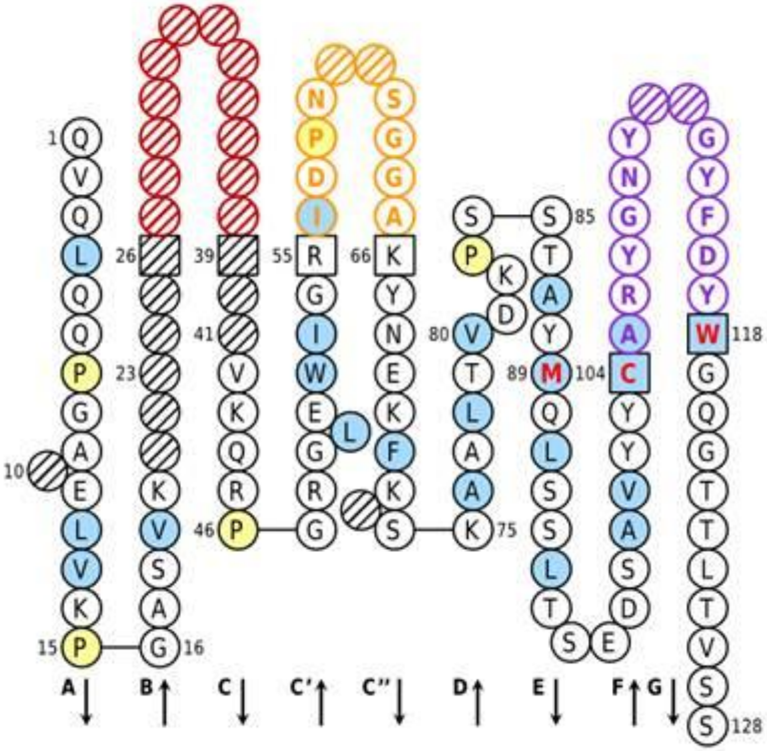
```
>seq2_TW
caggtccaactgcagcagcctggggctgagcttgtgaagcctggggcttcagtgaagctg
tgaagcagaggcctggacgaggccttgagtggattggaaggattgatcccaatagtggtg
gtgctaagtacaatgagaagttcaagagcaaggccgcactgactgtagacaaaccctcca
gcacagcctacatgcagctcagcagcctgacatctgaggactctgcggtctattattgtg
caagatacggtaactacggctactttgactactggggccaaggcaccactctcacagtct
cctcag
```

Result summary:	**Nucleotide deletions have been detected** (shown by dots in the alignments):			
	localization	nb of deleted nt	causing frameshift	from V-REGION codon	from nt position in user submitt
	FR1-IMGT	49	yes	21	60

IMGT/V-QUEST results after filling the deletion(s) gap(s): **Potentially productive IGH rearranged sequence** (no stop codon and in-frame junction) (Check also your sequence with BLAST against IMGT/GENE-DB reference sequences to eventually identify out-of-frame pseudogenes)			
V-GENE and allele	Musmus IGHV1-72*01 F	score = 1163	identity = **98.74%** (236/239 nt) [**98.33%** (235/239 nt)]
J-GENE and allele	Musmus IGHJ2*01 F	score = 231	identity = 97.92% (47/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Musmus IGHD2-1*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[19.14.38.11]	[X.8.11]	CARYGNYGYFDYW

7. V-REGION translation

```
                                  <------------------------------------------  FR1 - IMGT
                                  1               5                   10                  15
                                   Q   V   Q   L   Q   Q   P   G   A       E   L   V   K   P
seq2_TW                           cag gtc caa ctg cag cag cct ggg gct ... gag ctt gtg aag cct

AC163348 Musmus IGHV1-72*01 F     --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---

                                  ----------------------------------------> ______________
                                                  20                  25                  30
                                   G   A   S   V   K
seq2_TW                           ggg gct tca gtg aag ct. ... ... ... ... ... ... ... ... ...
                                                      L   S   C   K   A   S   G   Y   T   F
AC163348 Musmus IGHV1-72*01 F     --- --- --- --- --- --g tcc tgc aag gct tct ggc tac acc ttc

                                  __  CDR1 - IMGT  _____________ <-------------------------
                                                  35                  40                  45
                                                                          V   K   Q   R
seq2_TW                           ... ... ... ... ... ... ... ... ... ... ... gtg aag cag agg
                                                   T   S   Y   W   M   H   W
AC163348 Musmus IGHV1-72*01 F     ... ... ... ... acc agc tac tgg atg cac tgg --- --- --- ---

                                   FR2 - IMGT  ------------------------> _____________  CDR2
                                                  50                  55                  60
                                   P   G   R   G   L   E   W   I   G   R   I   D   P   N
seq2_TW                           cct gga cga ggc ctt gag tgg att gga agg att gat ccc aat ...

AC163348 Musmus IGHV1-72*01 F     --- --- --- --- --- --- --- --- --- --- --- --- --t --- ...

                                   - IMGT  __________ <---------------------------------------
                                                  65                  70                  75
                                       S   G   G   A   K   Y   N   E   K   F   K       S   K
seq2_TW                           ... agt ggt ggt gct aag tac aat gag aag ttc aag ... agc aag
                                                   T
AC163348 Musmus IGHV1-72*01 F     ... --- --- --- a-- --- --- --- --- --- --- --- ... --- ---

                                  ------------------------------  FR3 - IMGT  ---------------
                                                  80                  85                  90
                                   A   A   L   T   V   D   K   P   S   S   T   A   Y   M   Q
seq2_TW                           gcc gca ctg act gta gac aaa ccc tcc agc aca gcc tac atg cag
                                       T
AC163348 Musmus IGHV1-72*01 F     --- a-- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                  -------------------------------------------------------> ____
                                                  95                  100             104
                                   L   S   S   L   T   S   E   D   S   A   V   Y   Y   C   A
seq2_TW                           ctc agc agc ctg aca tct gag gac tct gcg gtc tat tat tgt gca

AC163348 Musmus IGHV1-72*01 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
```



A
B
C
C'
C''
D
E
F
G
1
10
15
16
23
26
39
41
46
55
66
75
80
85
89
104
118
128