6th Educational IgCLL Workshop – Uppsala 2016

Session II
Practical exercises

Vasilis Bikos

Session II - Practical exercises

- **An Introduction of IMGT / V-QUEST annotation tool**
    get familiar with the sequence submission process
    review basic parameters and options of the tool

- **Hands-on exercises**
    two sets of IGH and IGL rearrangement sequences
    annotate the sequences and comment on their features
    and/or their interpretation

- **Sequence interpretation and results**

Session II - Practical exercises

Our tool of choice : IMGT / V-QUEST

Access: **http://www.imgt.org/**

Select **"human"**

on the tab for **"immunoglobulin (IG) or antibody nucleotide sequences"**

Session II - Practical exercises - *Submit sequence(s)*

Analyse your IG or antibody nucleotide sequences

Your selection: Human

Your sequences will be compared to the **Human** *(Homo sapiens)* IG set from the IMGT/V-QUEST reference directory sets

Sequence sets to test IMGT/V-QUEST are available here

Sequence submission

◉ Type (or copy/paste) your nucleotide sequence(s) in FASTA format

```
>a
gcctcagtgaaagtctcctgcaaggcttctggatacaccttcacaaactaaactatcagctgggcatgacaggcccatggacaagggc
ttgaggaaatgggatggatcaacacgaacactgggaacctaacatatgcccagggcttcagaggacggtttgtcttctccatggacac
ctccgtcaccatgacatatcttcatatcagcagcctcaaggctgagaggcagtgttgtagtagtaccaggtactatacggtgactact
ggggccaggaaaccctggtcaccgtctcctcag
```

◯ Or give the path access to a local file containing your sequence(s) in FASTA format

[Choose File] No file chosen

[Start] [Clear the form]

Submit a single or a batch of sequences (up to 50) **:** *manual copy/paste or file upload* **:** *In FASTA format*

Advanced parameters

Selection of IMGT reference directory set	F+ORF+ in-frame P ▼		◉ With all alleles ◯ With allele *01 only	
Search for insertions and deletions in V-REGION	◉ Yes (slower, the nb of submitted sequences in a single run is limited to 10)		◯ No	
Parameters for IMGT/JunctionAnalysis	Nb of accepted D-GENE in IGH JUNCTION (default is 1)	default ▼	Nb of accepted mutations:	default ▼ in 3'V-REGION default ▼ in D-REGION default ▼ in 5'J-REGION
Parameters for "Detailed view"	Nb of nucleotides to exclude in 5' of the V-REGION for the evaluation of the nb of mutations (in results 9 and 10)		Nb of nucleotides to add (or exclude) in 3' of the V-REGION for the evaluation of the alignment score (in results 1)	

Session II - Practical exercises – *Retrieve results*

Display results

A. Detailed view ⦿ HTML ○ Text   Nb of nucleotides per line in alignments: [60 ▾]   Nb of aligned reference sequences: [5 ▾]

1. ☑ Alignment for V-GENE
2. ☐ Alignment for D-GENE
3. ☑ Alignment for J-GENE
4. ☑ Results of IMGT/JunctionAnalysis
   - ⦿ with full list of eligible D-GENE
   - ○ without list of eligible D-GENE
5. ☑ Sequence of the JUNCTION ('nt' and 'AA')

6. ☑ V-REGION alignment
7. ☑ V-REGION translation
8. ☑ V-REGION protein display
9. ☑ V-REGION mutation and AA change table
10. ☑ V-REGION mutation and AA change statistics
11. ☑ V-REGION mutation hotspots

12. ☑ Sequences of V-, V-J- or V-D-J- REGION ('nt' and 'AA') with gaps in FASTA and access to IMGT/PhyloGene for V-REGION ('nt')
13. ☑ Annotation by IMGT/Automat
14. IMGT Collier de Perles
   - ⦿ link to IMGT/Collier-de-Perles tool
   - ○ IMGT Collier de Perles (for a nb of sequences < 5)
   - ○ no IMGT Collier de Perles

Check all | Uncheck all | Default

B. Synthesis view ○ HTML ○ Text   Nb of nucleotides per line in alignments: [▾]   Summary table sequence order [▾]

1. ☑ Alignment for V-GENE
2. ☑ V-REGION alignment
3. ☑ V-REGION translation
4. ☑ V-REGION protein display

5. ☑ V-REGION protein display (with AA class colors)
6. ☑ V-REGION protein display (only AA changes displayed)
7. ☑ V-REGION most frequently occurring AA
8. ☑ Results of IMGT/JunctionAnalysis

Check all | Uncheck all | Default

C. Excel file ○ Open in a spreadsheet ○ Download in a zip archive ○ Display 1 CSV file in your browser

1. ☑ Summary
2. ☑ IMGT-gapped-nt-sequences
3. ☑ nt-sequences
4. ☑ IMGT-gapped-AA-sequences
5. ☑ AA-sequences
6. ☑ Junction

7. ☑ V-REGION-mutation-and-AA-change-table
8. ☑ V-REGION-nt-mutation-statistics
9. ☑ V-REGION-AA-change-statistics
10. ☑ V-REGION-mutation-hotspots
11. ☑ Parameters
12. ☑ scFv (only for option "Analysis of single chain Fragment variable (scFv)")

Check all | Uncheck all | Default

Information about the type of files provided for each option:
IMGT/V-QUEST Documentation
http://www.imgt.org/IMGT_vquest/share/textes/imgtvquest.html

Practical exercises – Solutions – Batch I

>Sequence A1

caggtgcagctacagcagtggggcgcaggactgttgaagccttcggagaccctgtc
cctcacctgcgctgtctatggtgggtccttcagtggttactactggagctggatccgcc
agcccccagggaaggggctggagtggattggggaaatcaatcatagtggaagcac
caactacaacccgtccctcaagagtcgagtcaccatatcagtagacacgtccaaga
accagttctccctaaagctgagctctgtgaccgccgcggacacggctgtgtattactg
tgcgagaggtctacctcttttggagtggttattgggtccttactactactactacggta
tggacgtctggggccaagggaccacggtcaccgtctcct

Sequence A1

Sequence: 1 A1

Sequence compared with the human IG set from the IMGT reference directory

>A1
caggtgcagctacagcagtggggcgcaggactgttgaagccttcggagaccctgtccctc
acctgcgctgtctatggtgggtccttcagtggttactactggagctggatccgccagccc
ccagggaaggggctggagtggattggggaaatcaatcatagtggaagcaccaactacaac
ccgtccctcaagagtcgagtcaccatatcagtagacacgtccaagaaccagttctcccta
aagctgagctctgtgaccgccgcggacacggctgtgtattactgtgcgagaggtctacct
cttttggagtggttattgggtccttactactactactacggtatggacgtctggggccaa
gggaccacggtcaccgtctcct

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV4-34*01 F	score = 1411	identity = 99.65% (284/285 nt)
J-GENE and allele	Homsap IGHJ6*02 F	score = 291	identity = 98.33% (59/60 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD3-3*01 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.10]	[8.7.22]	CARGLPLLEWLLGPYYYYYGMDVW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                    Score       Identity
AB019439 Homsap IGHV4-34*01 F       1411        99.65% (284/285 nt)
M99684 Homsap IGHV4-34*02 F         1402        99.30% (283/285 nt)
X92255 Homsap IGHV4-34*03 F         1402        99.30% (282/284 nt)
M95113 Homsap IGHV4-34*08 F         1402        99.30% (283/285 nt)
X56591 Homsap IGHV4-34*12 F         1402        99.30% (283/285 nt)
```

Sequence A1

```
                                <------------------------------------------------ FR1 - IMGT
                                1          5                 10                    15
A1                              cag gtg cag cta cag cag tgg ggc gca ... gga ctg ttg aag cct
AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---
M99684 Homsap IGHV4-34*02 F     --- --- --- --- --a --- --- --- --- ... --- --- --- --- ---
X92255 Homsap IGHV4-34*03 F     --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---
M95113 Homsap IGHV4-34*08 F     --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---
X56591 Homsap IGHV4-34*12 F     --- --- --- --- --- --- --- --- --- ... --- --- --- --- ---


                                ----------------------------------------------->  _____
                                          20                25                  30
A1                              tcg gag acc ctg tcc ctc acc tgc gct gtc tat ggt ggg tcc ttc
AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
M99684 Homsap IGHV4-34*02 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
X92255 Homsap IGHV4-34*03 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
M95113 Homsap IGHV4-34*08 F     --- --- --- --- --- --- --- --- --- --- --- --- --- a-- ---
X56591 Homsap IGHV4-34*12 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---


                                __  CDR1 - IMGT _____ <------------------------
                                          35                40                  45
A1                              ... ... ... ... agt ggt tac tac tgg agc tgg atc cgc cag ccc
AB019439 Homsap IGHV4-34*01 F   ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---
M99684 Homsap IGHV4-34*02 F     ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---
X92255 Homsap IGHV4-34*03 F     ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---
M95113 Homsap IGHV4-34*08 F     ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---
X56591 Homsap IGHV4-34*12 F     ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---


                                FR2 - IMGT ----------------------------> _____  CDR2
                                          50                55                  60
A1                              cca ggg aag ggg ctg gag tgg att ggg gaa atc aat cat agt ...
AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ...
M99684 Homsap IGHV4-34*02 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ...
X92255 Homsap IGHV4-34*03 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ...
M95113 Homsap IGHV4-34*08 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ...
X56591 Homsap IGHV4-34*12 F     --- --- --- --- --- --- --- --- --- --- --- --- -t- --- ...


                                - IMGT _____ <------------------------------------------
                                          65                70                  75
A1                              ... ... gga agc acc aac tac aac ccg tcc ctc aag ... agt cga
AB019439 Homsap IGHV4-34*01 F   ... ... --- --- --- --- --- --- --- --- --- ... --- ---
M99684 Homsap IGHV4-34*02 F     ... ... --- --- --- --- --- --- --- --- --- ... --- ---
X92255 Homsap IGHV4-34*03 F     ... ... --- --- --- --- --- --- --- --- --- ... --- ---
M95113 Homsap IGHV4-34*08 F     ... ... --- --- --- --- --- --- --- --- --- ... --- ---
X56591 Homsap IGHV4-34*12 F     ... ... --- --- --- --- --- --- --- --- --- ... --- ---


                                ------------------------------- FR3 - IMGT ----[----------
                                          80                85                  90
A1                              gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc cta aag
AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --g --
M99684 Homsap IGHV4-34*02 F     --- --- --- --- --- --- --- --- --- --- --- --- --g --
X92255 Homsap IGHV4-34*03 F     --- --- --- --- --- --- --- --- --- --- --- --- --g --
M95113 Homsap IGHV4-34*08 F     --- --- --- --- --- --- --- --- --- --- --- --- --g --
X56591 Homsap IGHV4-34*12 F     --- --- --- --- --- --- --- --- --- --- --- --- --g --


                                ------------------------------------------------[----->
                                          95                100                 104
A1                              ctg agc tct gtg acc gcc gcg gac acg gct gtg tat tac tgt gcg
AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
M99684 Homsap IGHV4-34*02 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
X92255 Homsap IGHV4-34*03 F     --- --- --- --- --- --- --- --- --- --c --- --- --- --
M95113 Homsap IGHV4-34*08 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
X56591 Homsap IGHV4-34*12 F     --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---


                                _____  CDR3 - IMGT _____
A1                              aga ggt cta cct ctt ttg gag tgg tta ttg ggt cct tac tac tac
AB019439 Homsap IGHV4-34*01 F   --- --
M99684 Homsap IGHV4-34*02 F     --- --
X92255 Homsap IGHV4-34*03 F
M95113 Homsap IGHV4-34*08 F
X56591 Homsap IGHV4-34*12 F     ---
```

Sequence A1

- **A single silent mutation**

- **VH CDR3 anchors and VH FR4 GXG motif are detected**

```
                                <------------------------------------- FR1 - IMGT
                                1           5                10              15
                                Q   V   Q   L   Q   Q   W   G   A       G   L   L   K   P
A1                              cag gtg cag cta cag cag tgg ggc gca ... gga ctg ttg aag cct

AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- ... --- --- --- --- ---


                                ------------------------------------------------->
                                            20              25              30
                                S   E   T   L   S   L   T   C   A   V   Y   G   G   S   F
A1                              tcg gag acc ctg tcc ctc acc tgc gct gtc tat ggt ggg tcc ttc

AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---


                                ___ CDR1 - IMGT _____    <-------------------------
                                                35              40              45
                                            S   G   Y   Y   W   S   W   I   R   Q   P
A1                              ... ... ... ... agt ggt tac tac tgg agc tgg atc cgc cag ccc

AB019439 Homsap IGHV4-34*01 F   ... ... ... ... --- --- --- --- --- --- --- --- --- --- ---


                                FR2 - IMGT --------------------------->          CDR2
                                            50              55              60
                                P   G   K   G   L   E   W   I   G   E   I   N   H   S
A1                              cca ggg aag ggg ctg gag tgg att ggg gaa atc aat cat agt ...

AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ...


                                - IMGT _____    <-----------------------------------
                                            65              70              75
                                        G   S   T   N   Y   N   P   S   L   K       S   R
A1                              ... ... gga agc acc aac tac aac ccg tcc ctc aag ... agt cga

AB019439 Homsap IGHV4-34*01 F   ... ... --- --- --- --- --- --- --- --- --- --- ... --- ---


                                ----------------------------------------- FR3 - IMGT -------------
                                            80              85                  90
                                V   T   I   S   V   D   T   S   K   N   Q   F   S   L   K
A1                              gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc cta aag

AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --g ---


                                -------------------------------------------------------------
                                            95              100             104
                                L   S   S   V   T   A   A   D   T   A   V   Y   Y   C   A
A1                              ctg agc tct gtg acc gcc gcg gac acg gct gtg tat tac tgt gcg

AB019439 Homsap IGHV4-34*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---


                                _____ CDR3 - IMGT _____
                                R   G   L   P   L   L   E   W   L   L   G   P   Y   Y   Y
A1                              aga ggt cta cct ctt ttg gag tgg tta ttg ggt cct tac tac tac

AB019439 Homsap IGHV4-34*01 F   --- --


                                _____
                                Y   Y   G   M   D   V   W   G   Q   G   T   T   V   T   V
A1                              tac tac ggt atg gac gtc tgg ggc caa ggg acc acg gtc acc gtc

AB019439 Homsap IGHV4-34*01 F


                                S
A1                              tcc t
```

Sequence A1

IGHV4-34*01

IGHD3-3*01

IGHJ6*02

identity 99,65%

in-frame

VH CDR3 length 22 aa

→ productive, unmutated IGH

Practical exercises – Solutions – Batch I

>Sequence A2

caggtgcagctggtgcagtctggagctgaagtgaagaagcctggggcctcagtgaa
ggtctcctgcaaggcttctggttacacctttacgaattatggtatcggctgggtgcga
caggcccctggacaagggcttgagtggatgggatggatcagcggttacaatggtga
tacaaactatgcacagaagttccaggacagagtcaccatgaccacagacacatcca
cgagcacagcctatatggacctggggagcctgagatctgacgacacggccgtgtat
tactgtgcgagagtagtggctacgctcccccccgtctactggggccagggaaccctg
gtcaccgtctcctca

Sequence A2

Sequence compared with the human IG set from the IMGT reference directory

>A2

```
caggtgcagctggtgcagtctggagctgaagtgaagaagcctggggcctcagtgaaggtc
tcctgcaaggcttctggttacacctttacgaattatggtatcggctgggtgcgacaggcc
cctggacaagggcttgagtggatgggatggatcagcggttacaatggtgatacaaactat
gcacagaagttccaggacagagtcaccatgaccacagacacatccacgagcacagcctat
atggacctgggggagcctgagatctgacgacacggccgtgtattactgtgcgagagtagtg
gctacgctcccccccgtctactggggccagggaaccctggtcaccgtctcctca
```

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV1-18*01 F	score = 1309	identity = 95.14% (274/288 nt)
J-GENE and allele	Homsap IGHJ4*02 F	score = 172	identity = 85.11% (40/47 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD5-12*01 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.11]	CARVVATLPPVYW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                  Score       Identity
M99641  Homsap IGHV1-18*01 F      1309        95.14% (274/288 nt)
HM855463 Homsap IGHV1-18*03 F     1300        94.79% (273/288 nt)
KC713938 Homsap IGHV1-18*04 F     1300        94.79% (273/288 nt)
X60503  Homsap IGHV1-18*02 (F)    1245        94.57% (261/276 nt)
M99637  Homsap IGHV1-8*01 F       1147        88.89% (256/288 nt)
```

Sequence A2

```
                          <------------------------------------------ FR1 - IMGT
                          1           5               10              15
A2                        cag gtg cag ctg gtg cag tct gga gct ... gaa gtg aag aag cct
M99641 Homsap IGHV1-18*01 F   --- --t --- --- --- --- --- --- --- .. --g --- --- --- ---
HM855463 Homsap IGHV1-18*03 F --- --t --- --- --- --- --- --- --- .. --g --- --- --- ---
KC713938 Homsap IGHV1-18*04 F --- --t --- --- --- --- --- --- --- .. --g --- --- --- ---
X60503 Homsap IGHV1-18*02 (F) --- --t --- --- --- --- --- --- --- .. --g --- --- --- ---
M99637 Homsap IGHV1-8*01 F    --- --- --- --- --- --- --- --g --- .. --g --- --- --- ---

                          -----------------------------------------> _____
                          20              25              30
A2                        ggg gcc tca gtg aag gtc tcc tgc aag gct tct ggt tac acc ttt
M99641 Homsap IGHV1-18*01 F   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
HM855463 Homsap IGHV1-18*03 F --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
KC713938 Homsap IGHV1-18*04 F --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
X60503 Homsap IGHV1-18*02 (F) --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
M99637 Homsap IGHV1-8*01 F    --- --- --- --- --- --- --- --- --- --- --- --a --- --- --c

                          ___ CDR1 - IMGT             <---------------------------
                                     35              40              45
A2                        ........ ... ... acg aat tat ggt atc ggc tgg gtg cga cag gcc
M99641 Homsap IGHV1-18*01 F   ... ... ... .. --c -gc --- --- --- a-- --- --- --- ---
HM855463 Homsap IGHV1-18*03 F ... ... ... .. --c -gc --- --- --- a-- --- --- --- ---
KC713938 Homsap IGHV1-18*04 F ... ... ... .. --c -gc --c --- --- a-- --- --- --- ---
X60503 Homsap IGHV1-18*02 (F) ... ... ... .. --c -gc --- --- --- a-- --- --- --- ---
M99637 Homsap IGHV1-8*01 F    ... ... ... .. --c -g- --- -a- --- aa- -- --- --- ---

                          FR2 - IMGT  -----------------------------> _____ CDR2
                                     50              55              60
A2                        cct gga caa ggg ctt gag tgg atg gga tgg atc agc ggt tac ...
M99641 Homsap IGHV1-18*01 F   --- --- --- --- --- --- --- --- --- --- --- -c- --- ...
HM855463 Homsap IGHV1-18*03 F --- --- --- --- --- --- --- --- --- --- --- -c- --- ...
KC713938 Homsap IGHV1-18*04 F --- --- --- --- --- --- --- --- --- --- --- -c- --- ...
X60503 Homsap IGHV1-18*02 (F) --- --- --- --- --- --- --- --- --- --- --- -c- --- ...
M99637 Homsap IGHV1-8*01 F    a-- --- --- --- --- --- --- --- --- --g -a- cc- a-- ...

                          - IMGT _____ <-------------------------------------
                                     65              70              75
A2                        .. aat ggt gat aca aac tat gca cag aag ttc cag ... gac aga
M99641 Homsap IGHV1-18*01 F   .. --- --- a-c --- --- --- --- --- --- c-- --- ... -g- ---
HM855463 Homsap IGHV1-18*03 F .. --- --- a-c --- --- --- --- --- --- c-- --- ... -g- ---
KC713938 Homsap IGHV1-18*04 F .. --- --- a-c --- --- --- --- --- --- c-- --- ... -g- ---
X60503 Homsap IGHV1-18*02 (F) .. --- --- a-c --- --- --- --- --- --- c-- --- ... -g- ---
M99637 Homsap IGHV1-8*01 F    .. -g- --- a-c --- gg- --- --- --- --- --- --- ... -g- ---

                          -------------------------------- FR3 - IMGT --------------
                                     80              85              90
A2                        gtc acc atg acc aca gac aca tcc acg agc aca gcc tat atg gac
M99641 Homsap IGHV1-18*01 F   --- --- --- --- --- --- --- --- --- --- --- --c --- --g
HM855463 Homsap IGHV1-18*03 F --- --- --- --- --- --- --- --- --- --- --- --c --- --g
KC713938 Homsap IGHV1-18*04 F --- --- --- --- --- --- --- --- --- --- --- --c --- --g
X60503 Homsap IGHV1-18*02 (F) --- --- --- --- --- --- --- --- --- --- --- --c --- --g
M99637 Homsap IGHV1-8*01 F    --- --- --- --- -gg a-- --c --- -ta --- --- --- --c --- --g

                          ------------------------------------------------------> ____
                                     95              100             104
A2                        ctg ggg agc ctg aga tct gac gac acg gcc gtg tat tac tgt gcg
M99641 Homsap IGHV1-18*01 F   --- a-- --- --- --- --- --- --- --- --- --- --- --- ---
HM855463 Homsap IGHV1-18*03 F --- a-- --- --- --- --- --- --- -t- --- --- --- --- ---
KC713938 Homsap IGHV1-18*04 F --- a-- --- --- --- --- --- --- --- --- --- --- --- ---
X60503 Homsap IGHV1-18*02 (F) --- a-- --- --a --- --- --- --- --- ---
M99637 Homsap IGHV1-8*01 F    --- a-c --- --- --- --- --g --- --- --- --- --- --- ---

                          _____ CDR3 - IMGT _____
A2                        aga gta gtg gct acg ctc ccc ccc gtc tac tgg ggc cag gga acc
M99641 Homsap IGHV1-18*01 F   --- -a
HM855463 Homsap IGHV1-18*03 F --- -a
KC713938 Homsap IGHV1-18*04 F --- -a
X60503 Homsap IGHV1-18*02 (F)
M99637 Homsap IGHV1-8*01 F    --- -g
```

Sequence A2

9. V-REGION mutation and AA change table

FR1-IMGT	CDR1-IMGT	FR2-IMGT	CDR2-IMGT	FR3-IMGT	CDR3-IMGT
t6>g g33>a	c105>g g107>a, S36>N (- - -) c108>t, S36>N (- - -)	a118>g, S40>G (+ + -)	c173>g, A58>G (- + -) a190>g, N64>D (+ + -) c192>t, N64>D (+ + -)	c211>t, L71>F (+ - -) g221>a, G74>D (- - -) c264>t g270>c, E90>D (+ - +) a274>g, R92>G (- - -)	a320>t

http://www.imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html

Amino acid classification

- **Hydropathy**
- **Volume**
- **Chemical**
- **Charge**
- **Hydrogen donor/acceptor**
- **Polarity**

IMGT 'Physicochemical' classes of the 20 common amino acids ★

Eleven IMGT 'Physicochemical' classes of the 20 common amino acids have been defined by the physicochemical properties of their side chains (Figure 1) [1]. These standardized classes are used in IMGT® databases and tools, for the description of amino acid class changes in mutations and protein engineering.

Figure 1: The 11 IMGT 'Physicochemical' classes of the 20 common amino acids [1].

Sequence A2

IGHV1-18*01
IGHD5-12*01
IGHJ4*02
identity 95,14%
in-frame
VH CDR3 length 11 aa

→ productive, mutated IGH

Practical exercises – Solutions – Batch I

>Sequence A4

caggtgcagctgcaggagtcgggcccaggactggtgaagccttcggggaccctgtc
cctcacctgcgctgtctctggtggctccatcagcagtactaactggtggagttgggtc
cgccagcccccagggaaggggctggagtggattggagaaatctatcatagtggga
gcaccaactacaacccgtccctcaagagtcgagtcaccatatcaatagacaagtcc
aagaaccagttctcgctgaatctgaggtcagtgaccgccgcggacacggccgtgtat
tactgtgcgagaggttgtagtggtgggagctgccaatttgactactggggccaggga
accctggtcaccgtctcc

Sequence A4

Sequence: 1 A4

Sequence compared with the human IG set from the IMGT reference directory

>A4

gaggtgcagctggtggagtctggggggaggcctggtcaagcctggggggtccctgagactc
tcctgtgcagcctctggattcaccttcagtagctatagtatgaactgggtccgccaggct
ccaggaaggggctggagtgggtctcatccatgagcagtagtactagttacaaaaactac
gcagactcagtgaagggccgattcaccatctccagagacaacgcccagaactcactgtat
ctgcaaatgaacagcctgagagccgaggacacggctgtgtattactgtgcgagagacgcg
aatggtatggacgtctggggccaagggaccacggtcaccgtctcctca

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV3-21*01 F	score = 1372	identity = 97.57% (281/288 nt)
J-GENE and allele	Homsap IGHJ6*02 F	score = 215	identity = 90.38% (47/52 nt)
D-GENE and allele by IMGT/JunctionAnalysis	No results		-
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.9]	CARDANGMDVW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

	Score	Identity
AB019439 Homsap IGHV3-21*01 F	1372	97.57% (281/288 nt)
M99658 Homsap IGHV3-21*02 F	1363	97.22% (280/288 nt)
HM855323 Homsap IGHV3-21*03 F	1363	97.22% (280/288 nt)
HM855688 Homsap IGHV3-21*04 F	1363	97.22% (280/288 nt)
HM855336 Homsap IGHV3-48*04 F	1300	94.79% (273/288 nt)

Sequence A4

No D-gene has been identified. But the sequence itself is more important than gene names

4. Results of IMGT/JunctionAnalysis

Maximum number of accepted mutations in: 3'V-REGION = 2, D-REGION = 4, 5'J-REGION = 2
Maximum number of accepted D-GENE = 1

Analysis of the JUNCTION

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N	5'J-REGION	J name	Vmut	Jmut	Ngc
A4	Homsap IGHV3-21*01	tgtgcgagaga	cgcgaat	...................ggtatggacgtctgg	Homsap IGHJ6*02	0	0	4/7

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	A	R	D	A	N	G	M	D	V	W					
A4	tgt	gcg	agg	gac	gcg	aat	ggt	atg	gac	gtc	tgg	+	9	1,237.38	4.44	CARDANGMDVW

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

Subset #2 features

- IGHV3, mostly IGHV3-21
- Mostly IGHJ6
- Mostly mutated
- VH CDR3 : 9 aa
- Mostly D or E at position VH CDR3-107

This rearrangement represents a CLL subset#2 case. Thus it is related with poor prognosis and this should be stated clearly in the report

Sequence A4

IGHV3-21*01
IGHD ND
IGHJ6*02
identity 97,57%
in-frame
VH CDR3 length 9 aa

→ productive, borderline mutated IGH, subset #2 case

Practical exercises – Solutions – Batch I

>Sequence A5

gaggtgcagctggtggagtctggggggaggcttggtccagcctggggggtccctgag
actctcctgtgcagcctctggattcaccttcagtagctatgctatgcactgggtccgcc
aggctccagggaagggactggaatatgtttcagctattagtagtaatggggggtagc
acatattatgcaaactctgtgaagggcagattcaccatctccagagacaattccaag
aacacgctgtatcttcaaatgggcagcctgagagctgaggacatggctgtgtattac
tgtgcgagcgctgggagtccttctatggtggccacgagaggactactactacta
catggacgtctggggccaagggaccctggtcac

Sequence A5

Sequence compared with the human IG set from the IMGT reference directory

>A5

```
gaggtgcagctggtggagtctgggggaggcttggtccagcctggggggtccctgagactc
tcctgtgcagcctctggattcacctttcagtagctatgctatgcactgggtccgccaggct
ccagggaagggactggaatatgtttcagctattagtagtaatgggggtagcacatattat
gcaaactctgtgaagggcagattcaccatctccagagacaattccaagaacacgctgtat
cttcaaatgggcagcctgagagctgaggacatggctgtgtattactgtgcgagcgctggg
agtccttctatggtggccacgagaggactactactactacatggacgtctggggcca
agggaccctggtcac
```

Result summary:	Unproductive IGH rearranged sequence (out-of-frame junction)		
V-GENE and allele	Homsap IGHV3-64*01 F	score = 1435	identity = 100.00% (288/288 nt)
J-GENE and allele	Homsap IGHJ6*03 F	score = 215	identity = 90.38% (47/52 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD5-12*01 F	D-REGION is in reading frame 2	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.7]	[8.8.X]	CASAGSPSMVATR##YYYYYMDVW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                    Score      Identity
M99682 Homsap IGHV3-64*01 F         1435       100.00% (288/288 nt)
AB019437 Homsap IGHV3-64*02 F       1417       99.31% (286/288 nt)
KC713941 Homsap IGHV3-64D*06 F      1363       97.22% (280/288 nt)
M77301 Homsap IGHV3-64*05 F         1354       96.88% (279/288 nt)
M77298 Homsap IGHV3-64*03 F         1345       96.53% (278/288 nt)
```

Sequence A5

from IMGT scientific chart:

http://www.imgt.org/IMGTScientificChart/Numbering/IMGTnumbering.html

IMGT unique numbering for all IG and TR V-REGIONs of all species: interest for structure and evolution

Citing IMGT unique numbering for V-REGION: Lefranc, M.-P. et al., Dev. Comp. Immunol., 27, 55-77 (2003) PMID: 12477501

In order to easily compare V-REGION sequences of IG and TR from all species, a unique numbering has been defined by Marie-Paule Lefranc [1-6].

The 'IMGT unique numbering' concept is part of the 'NUMEROTATION' concept of IMGT-ONTOLOGY [7]

The IMGT unique numbering for all IG and TR V-REGIONs of all species relies on the high conservation of the structure of the variable region [1-6]. This numbering, set up after aligning more than 5 000 sequences, takes into account and combines the definition of the framework (FR) and complementarity determining regions (CDR) [8], structural data from X-ray diffraction studies [9], and the characterization of the hypervariable loops [10]. The delimitations of the FR-IMGT and CDR-IMGT regions have been defined and Correspondence between the IMGT unique numbering and the other numberings has been established [6].

The IMGT unique numbering has many advantages:

- It allows an easy comparison between sequences coding the variable regions, whatever the antigen receptor (IG or TR), the chain type (heavy or light chains for IG [4, 11]; alpha, beta, gamma or delta chains for TR [5]), or the species.
- In the IMGT unique numbering, the conserved amino acids always have the same position, for instance Cystein 23 (1st-CYS), Tryptophane 41 (CONSERVED-TRP), Leucine 89, Cystein 104 (2nd-CYS). The hydrophobic amino acids of the framework regions are also found in conserved positions. IMGT Colliers de perles of the human germline immunoglobulin IGLV6-57 gene and T cell receptor TRGV9 gene can be chosen as examples.
- This unique numbering has allowed the redefinition of the limits of the FR and CDR. The FR-IMGT and CDR-IMGT lengths themselves become crucial information characterizing the variable regions belonging to a group, a subgroup and/or a gene [4,5].
- Framework amino acids (and codons) located at the same position in different sequences can be compared without requiring sequence alignments [11,12]. This also holds for amino acids belonging to CDR-IMGT of the same length.
- The IMGT unique numbering has allowed a standardized IMGT description of mutations for the IMGT description of allele polymorphisms and somatic hypermutations of the variable regions. The mutations and allelic polymorphisms are described by comparison to the germline V-REGION, D-REGION, J-REGION and to the C-REGION from the IMGT reference directory. Based on these criteria, Alignments of alleles [4,5] and Tables of alleles have been set up.
- The IMGT unique numbering has allowed to set up the IMGT standardized criteria for the statistical analysis of IG and TR V-REGION amino acid porperties [11].
- The unique numbering is used as the output of IMGT/V-QUEST integrated sequence alignment tool which analyzes your IG and TR variable (germline or rearranged) sequences according to IMGT criteria. In IMGT/V-QUEST, a variable rearranged sequence is compared to the appropriate sets of V-REGION, D-REGION and J-REGION alleles from the IMGT reference directory. The results show, aligned with the input sequence, the sequences of the most homologous V-REGION alleles and, if appropriate, D-REGION (for heavy chains) and J-REGION allele. The aligned V-REGION sequences are displayed according to the IMGT unique numbering and with the FR-IMGT and CDR-IMGT delimitations.

By facilitating comparisons between the sequences and the descriptions of alleles and mutations, the IMGT unique numbering represents a big step forward in the analysis of the IG and TR sequences of all species. Moreover, it gives in-sight in the structural configuration of the variable domain and opens interesting views on the evolution of the sequences of the V-set, since this numbering has been applied with success to all the sequences belonging to the V-set of the immunoglobulin superfamily, including nonrearranging sequences in vertebrates (CD4, Xenopus CTX,...) and in invertebrates (Drosophila Amalgam, Drosophila Fasciclin II, etc.) [1, 2, 9].

In the IMGT unique numbering, the conserved amino acids always have the same position, for instance **Cystein 23 (1st-CYS), Tryptophane 41 (CONSERVED-TRP), Leucine 89, Cystein 104 (2nd-CYS).** The hydrophobic amino acids of the framework regions are also found in conserved positions.

Sequence A5

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	110	111	111.1	111.2	111.3	111.4	112.5	112.4	112.3	112.2	112.1	112	113	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	A	S	A	G	S	P	S	N	V	A	T	R	#	#	Y	Y	Y	Y	Y	M	D	V	W					
A5	tgt	gcg	agc	gct	ggg	agt	cct	tct	atg	gtg	gcc	acg	aga	gg-	.ac	tac	tac	tac	tac	tac	atg	gac	gtc	tgg	-	NR	NR	NR	NR

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

What to do when this is the single rearrangement amplified?

The amplification reaction should be repeated:
using different sets of primers in order to
      (i) confirm the presence of the unproductive rearrangement
      (ii) look for the presence of a second productive rearrangement.

Sequence A5

IGHV3-64*01
IGHD5-12*01
IGHJ6*03
identity 100,00%
out-of-frame
VH CDR3 length X

→ **unproductive, unmutated IGH**

Practical exercises – Solutions – Batch I

>**Sequence A7**

gaggtgcagctggtggagtctgggtgaggcttggtacagcctggagggtccctgag
actctcctgtgcagcctctggattcaccttcagtagctcctggatgcactgggtctgcc
aggctccggagaaggggctggagtgggtggccgacataaagtgtgacggaagtga
caaatactatgtagactctgtgaagggccgattgaccatctccagagacaatgccaa
gaactccctctatctgcaagtgaacagcctgagagctgaggacatgaccgtgtatta
ctgctgggggtcttacgatattttgactggttattacgccgggaccgctcacaactgg
ttcgacccctggggccagggaaccctggtcaccgtctcct

Sequence A7

Sequence: 1 A7

Sequence compared with the human IG set from the IMGT reference directory

>A7
gaggtgcagctggtggagtctggggtgaggcttggtacagcctggagggtccctgagactc
tcctgtgcagcctctggattcacctttcagtagctcctggatgcactgggtctgccaggct
ccggagaaggggctggagtgggtggccgacataaagtgtgacggaagtgacaaatactat
gtagactctgtgaagggccgattgaccatctccagagacaatgccaagaactccctctat
ctgcaagtgaacagcctgagagctgaggacatgaccgtgtattactgctgggggtcttac
gatattttgactggttattacgccgggaccgctcacaactggttcgacccctggggccag
ggaaccctggtcaccgtctcct

Result summary:	Unproductive IGH rearranged sequence (stop codons)		
V-GENE and allele	Homsap IGHV3-52*01 P	score = 1426	identity = 99.31% (286/288 nt)
J-GENE and allele	Homsap IGHJ5*02 F	score = 240	identity = 100.00% (48/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD3-9*01 F	D-REGION is in reading frame 2	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.10]	[8.8.21]	CWGSYDILTGYYAGTAHNWFDPW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                 Score       Identity
M99678  Homsap IGHV3-52*01 P     1426        99.31% (286/288 nt)
Z17388  Homsap IGHV3-52*02 P     1417        98.96% (285/288 nt)
J00237  Homsap IGHV3-52*03 P     1417        98.96% (285/288 nt)
HM855666 Homsap IGHV3-7*03 F     1201        90.62% (261/288 nt)
M99649  Homsap IGHV3-7*01 F      1192        90.28% (260/288 nt)
```

Sequence A7

from IMGT scientific chart discussing functionality:

http://www.imgt.org/IMGTScientificChart/SequenceDescription/IMGTfunctionality.html

Functionality

The definition of functionality is based on the sequence analysis [1].

The 'Functionality' concept is part of the 'IDENTIFICATION' axiom of IMGT-ONTOLOGY[2]

For IG and TR "Germline" entities and for "C-GENE"

A germline entity (V-GENE, D-GENE or J-GENE) or a C-GENE can be functional, ORF or pseudogene.

FUNCTIONAL
A germline entity (V-GENE, D-GENE or J-GENE) or a C-GENE is functional if the coding region has an open reading frame without stop codon, and if there is no described defect in the splicing sites, recombination signals and/or regulatory elements.

ORF (Open Reading Frame)
A germline entity (V-GENE, D-GENE or J-GENE) or a C-GENE is qualified as ORF (Open Reading Frame) if the coding region has an open reading frame, but:

- alterations have been described in the splicing sites, recombination signals and/or regulatory elements.
- and/or changes of conserved amino acids have been suggested by the authors to lead to uncorrect folding.
- and/or the entity is an orphon.

PSEUDOGENE
A germline entity (V-GENE, D-GENE or J-GENE) or a C-GENE is qualified as pseudogene if the coding region has stop codon(s) and/or frameshift mutation(s).
In particular, a V-GENE is considered as pseudogene if these defects occur in the L-PART1 and/or V-EXON, or if there is a mutation in the L-PART1 INIT-CODON atg.
A J-GENE is considered as pseudogene if it has been identified by the presence of a recombination signal upstream of an open reading frame, but it has no donor splicing site in 5' or the donor splice is not in the expected sf1 or if it has no conserved Phe(or Trp)-Gly-X-Gly motif. If the defects are important, pseudogenes can eventually also be qualified as vestigial (vg), for example: a germline V-GENE which cannot be assigned to a given subgroup because it is too divergent from the other pseudogenes and has too many stop codons and frameshifts.
Only vestigial pseudogenes reported in Gene tables are qualified as pseudogenes in the List of IG and TR genes and in the Potential germline repertoires.

A V-gene is considered as a pseudogene if:

- there are defects along its Leader region or the V-region

- there is a mutation at the initial codon of the Leader-Part1 region

Sequence A7

IGHV3-52*01 is a pseudogene because it encodes for a **stop codon** at **VH FR1-9** position

Sequence A7

IGHV3-52*01 P
IGHD3-9*01
IGHJ5*02
identity 99,31%
in-frame
VH CDR3 length 21

→ **unmutated IGH, unproductive due to IGHV pseudogene utilization**

Practical exercises – Solutions – Batch I

>Sequence A9

gacatccagatgacccagtctccatcctccctgtctgcatctgtaggagacagagtca
ccatcacttgccaggcgagtcaggacattagcaactatttaaattggtatcagcaga
aaccagggaaagcccctaagctcctgatctacgatgcatccaatttggaaacaggg
gtcccatcaaggttcagtggaagtggatctgggacagattttactttcaccatcagca
gcctgcagcctgaagatattgcaacatattactgtcaacagcttggtggttccccgtg
gaccttcggccaagggaccaagctggaaatcaaacgta

Sequence A9

Sequence compared with the human IG set from the IMGT reference directory

>A9
gacatccagatgacccagtctccatcctccctgtctgcatctgtaggagacagagtcacc
atcacttgccaggcgagtcaggacattagcaactatttaaattggtatcagcagaaacca
gggaaagcccctaagctcctgatctacgatgcatccaatttggaaacagggggtcccatca
aggttcagtggaagtggatctgggacagattttactttcaccatcagcagcctgcagcct
gaagatattgcaacatattactgtcaacagcttggtggttccccgtggaccttcggccaa
gggaccaagctggaaatcaaacgta

Result summary:	Productive IGK rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGKV1-33*01 F, or Homsap IGKV1D-33*01 F	score = 1345	identity = 98.21% (274/279 nt)
J-GENE and allele	Homsap IGKJ1*01 F	score = 172	identity = 94.74% (36/38 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[26.17.36.10]	[6.3.9]	CQQLGGSPWTF

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon plus 15 nt of the CDR3-IMGT)

		Score	Identity
M64856	Homsap IGKV1-33*01 F	1345	98.21% (274/279 nt)
M64855	Homsap IGKV1D-33*01 F	1345	98.21% (274/279 nt)
X59315	Homsap IGKV1-39*01 F	1165	91.04% (254/279 nt)
X59312	Homsap IGKV1D-39*01 F	1165	91.04% (254/279 nt)
X63398	Homsap IGKV1-27*01 F	1147	90.32% (252/279 nt)

Sequence A9

The sequence is aligned to both IGKV1-33 and IGKV1D-33
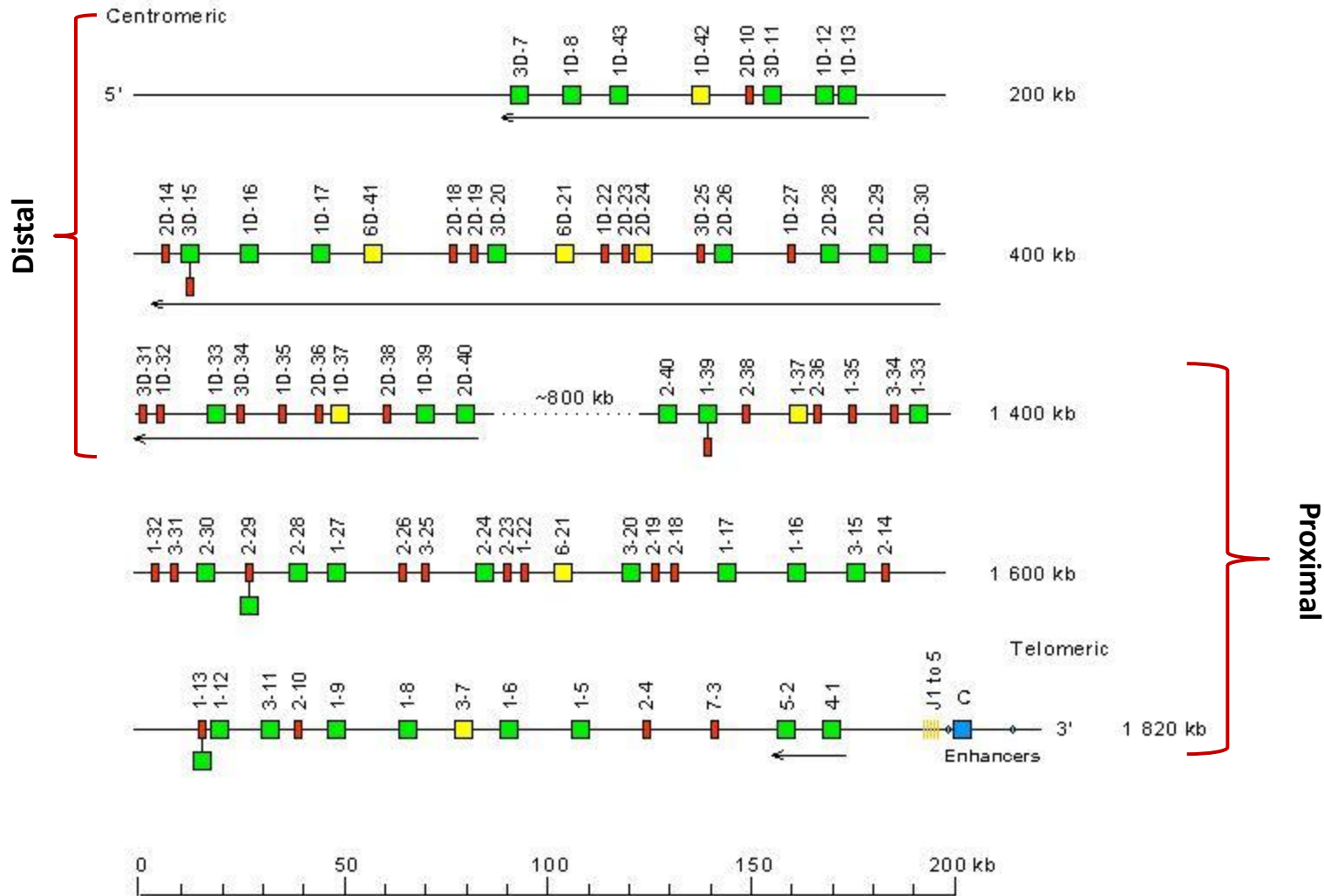
Distal V-CLUSTER: Duplication of proximal V-CLUSTER

Sequence A9

- All the identified mutations are **clustered at the end of the rearrangement**, along the VK CDR3 region

- (%) percentage of identity is **evaluated** from the 1st IGKV nucleotide to the 2nd-CYS codon **plus 15nt of the VK CDR3**

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon plus 15 nt of the CDR3-IMGT)

			Score	Identity
M64856	Homsap IGKV1-33*01 F		1345	98.21% (274/279 nt)
M64855	Homsap IGKV1D-33*01 F		1345	98.21% (274/279 nt)
X59315	Homsap IGKV1-39*01 F		1165	91.04% (254/279 nt)
X59312	Homsap IGKV1D-39*01 F		1165	91.04% (254/279 nt)
X63398	Homsap IGKV1-27*01 F		1147	90.32% (252/279 nt)

Alignment with FR-IMGT and CDR-IMGT delimitations

Sequence A9

Maximum number of accepted mutations in: 3'V-REGION = 2, 5'J-REGION = 2
Maximum number of accepted D-GENE = 0

Analysis of the JUNCTION

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N	P	5'J-REGION	J name	Vmut	Jmut	Ngc
A9	Homsap IGKV1-33*01	tgtcaacagcttg.............	gtggttccc	c	gtggaccttc	Homsap IGKJ1*01	2	1	6/9

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	Q	Q	L	G	G	S	P	W	T	F					
A9	tgt	caa	cag	ctt	ggt	ggt	tcc	ccg	tgg	acc	ttc	+	9	1,223.38	5.49	CQQLGGSPWTF

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

- case of **extreme exonuclease activity** in the 3' end of the V gene

- after 2nd-CYS codon **only 10 nucleotides belong to the V gene**

Sequence A9

Advanced parameters

Selection of IMGT reference directory set	F+ORF+ in-frame P ▼	◉ With all alleles ○ With allele *01 only
Search for insertions and deletions in V-REGION	○ Yes (slower, the nb of submitted sequences in a single run is limited to 10)	◉ No

Parameters for IMGT/JunctionAnalysis — Nb of accepted D-GENE in IGH JUNCTION (default is 1) — default ▼ — Nb of accepted mutations: default ▼ in 3'V-REGION / default ▼ in D-REGION / default ▼ in 5'J-REGION

Parameters for "Detailed view" — Nb of nucleotides to exclude in 5' of the V-REGION for the evaluation of the nb of mutations (in results 9 and 10) — Nb of nucleotides to add (or exclude) in 3' of the V-REGION for the evaluation of the alignment score (in results 1) — −5

Sequence: 1 A9

Sequence compared with the human IG set from the IMGT reference directory

```
>A9
gacatccagatgacccagtctccatcctccctgtctgcatctgtaggagacagagtcacc
atcacttgccaggcgagtcaggacattagcaactatttaaattggtatcagcagaaacca
gggaaagcccctaagctcctgatctacgatgcatccaatttggaaacaggggtcccatca
aggttcagtggaagtggatctgggacagatttttactttcaccatcagcagcctgcagcct
gaagatattgcaacatattactgtcaacagcttggtggttccccgtggaccttcggccaa
gggaccaagctggaaatcaaacgta
```

Result summary:	Productive IGK rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGKV1-33*01 F, or Homsap IGKV1D-33*01 F	score = 1347	identity = 99.27% (272/274 nt)
J-GENE and allele	Homsap IGKJ1*01 F	score = 172	identity = 94.74% (36/38 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[26.17.36.10]		[6.3.9]	CQQLGGSPWTF

---

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon plus 10 nt of the CDR3-IMGT)

		Score	Identity
M64856	Homsap IGKV1-33*01 F	1347	99.27% (272/274 nt)
M64855	Homsap IGKV1D-33*01 F	1347	99.27% (272/274 nt)
X59315	Homsap IGKV1-39*01 F	1176	91.97% (252/274 nt)
X59312	Homsap IGKV1D-39*01 F	1176	91.97% (252/274 nt)
X63398	Homsap IGKV1-27*01 F	1158	91.24% (250/274 nt)

Practical exercises – Solutions – Batch II

>Sequence B1

gaggaggaggtgttggagtctgggggaggcttggtacagcctggggggtccctgag
actctcctgtgtagcctctggattcacctttagtaattatgtcatgagttgggtccgcc
aggctccagggaggggactggagtgggtctccgttattactgatagtggtggtcata
catactatgcagactccgtgaagggccggttcaccatctccagagacaattccaaga
acaccctatatcttcaaatgaacagcctgagagtcgaggacacggccagatattact
gtgcgagagactggggggatattgtgggggaggcaccctggtcaccgtctccaca
t

Sequence B1

Sequence compared with the human IG set from the IMGT reference directory

>B1

```
ccatctccggcgacagtgtctctaccaacagtgctgcttggaactggatcaggcagtccc
catcgagaggccttgagtggctgggaaggacatactacaggaactccaactggttaaatg
attatgcagtttctgtcaaaggtcgaatagtcatctatccagacacatccaagaaccagt
tctccctggacctgaactctgtgactcccgaggacacggctgtgtattactgtgcaagag
gcgttgggtgggagattccagaggcccatggctgtgactactggggccagggaaccctgg
tcaccgtcttcctcaggta
```

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction) (a)		
V-GENE and allele	Homsap IGHV6-1*01 F, or Homsap IGHV6-1*02 F	score = 341	identity = 60.94% (142/233 nt)
J-GENE and allele	Homsap IGHJ4*02 F	score = 159	identity = 81.25% (39/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD2-15*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[3.17.38.11]	[10.9.16]	CARGVGWEIPEAHGCDYW

(a) Low V-REGION identity (60.94%): this may indicate potential nucleotide insertion(s) and/or deletion(s). Try 'Search for insertions and deletions' in 'Advanced parameters' at the bottom of the Search page

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                               Score    Identity
J04097 Homsap IGHV6-1*01 F     341      60.94% (142/233 nt)
Z14223 Homsap IGHV6-1*02 F     341      60.94% (142/233 nt)
Z75355 Homsap IGHV4-4*06 F     178      53.24% (115/216 nt)
Z75356 Homsap IGHV4-34*13 F    160      53.05% (113/213 nt)
Z75359 Homsap IGHV4-59*09 F    158      52.11% (111/213 nt)
```

Warning message!!

Sequence B1

An overall bad alignment

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                              Score    Identity
J04097 Homsap IGHV6-1*01 F    341      60.94% (142/233 nt)
Z14223 Homsap IGHV6-1*02 F    341      60.94% (142/233 nt)
Z75355 Homsap IGHV4-4*06 F    178      53.24% (115/216 nt)
Z75356 Homsap IGHV4-34*13 F   160      53.05% (113/213 nt)
Z75359 Homsap IGHV4-59*09 F   158      52.11% (111/213 nt)
```

Alignment with FR-IMGT and CDR-IMGT delimitations

```
                              <---------------------------------- FR1-IMGT ---------------
B1                            ..........................................................
J04097 Homsap IGHV6-1*01 F    caggtacagctgcagcagtcaggtcca...ggactggtgaagccctcgcagaccctctca
Z14223 Homsap IGHV6-1*02 F    caggtacagctgcagcagtcaggtccg...ggactggtgaagccctcgcagaccctctca
Z75355 Homsap IGHV4-4*06 F    ..........................................................
Z75356 Homsap IGHV4-34*13 F   ..........................................................
Z75359 Homsap IGHV4-59*09 F   ..........................................................

                              ----------------->_____ CDR1-IMGT _____<-----
B1                            .........ccatctccggcgacagtgtctctacc......aacagtgctgcttggaactgg
J04097 Homsap IGHV6-1*01 F    ctcacctgtgc-at-tc--gggacagtgtct-t.......-g--acag----gcttggaac
Z14223 Homsap IGHV6-1*02 F    ctcacctgtgc-at-tc--gggacagtgtct-t.......-g--acag----gcttggaac
Z75355 Homsap IGHV4-4*06 F    ..................tct-gtg-ctc-atc-g-...........---ag-aac---tgga-t
Z75356 Homsap IGHV4-34*13 F   ..................tat-gtg-gtc--tc...........----g-tac-actgga-c
Z75359 Homsap IGHV4-59*09 F   ..................tct-gtg-ctc-atc...........----ag-tac-actgga-c

                              --------------- FR2-IMGT -------------------->_____ CDR
B1                            atcaggcagtccccatcgagaggccttgagtggctgggaaggacatactacaggaactcc
J04097 Homsap IGHV6-1*01 F    tgg-tcag-cagt-cc-atcga-aggccttga-tg-ctgg-a-ggaca---tac-gg---
Z14223 Homsap IGHV6-1*02 F    tgg-tcag-cagt-cc-atcga-aggccttga-tg-ctgg-a-ggaca---tac-gg---
Z75355 Homsap IGHV4-4*06 F    tgggtc-gccag--cc-ag-gannnggct-ga-tg-attg--ga-at---tcat-gt...
Z75356 Homsap IGHV4-34*13 F   tgg-tc-gccag--cc-ag-gaaggggct-ga-tg-attg--ga-at-a-tcat-gt...
Z75359 Homsap IGHV4-59*09 F   tgg-tc-g-cag--cc-ag-nannngact-ga-tg-attg--tatat---ttac-gt...

                              2-IMGT _____<------------------------------------------
B1                            ...aactggttaaatgattatgcagtttctgtcaaa...ggtcgaatagtcatctatcca
J04097 Homsap IGHV6-1*01 F    ...--g----at--------------a-----g---...a--------ac---a-c----
Z14223 Homsap IGHV6-1*02 F    ...--g----at--------------a-----g---...a--------ac---a-c----
Z75355 Homsap IGHV4-4*06 F    ......g--agc-cca-c--caacccg--cc----g...a----g-cac---g-cagt-
Z75356 Homsap IGHV4-34*13 F   ......g-aagc-cca-c--caacccc--cc----g...a----g-cac---a-cagt-
Z75359 Homsap IGHV4-59*09 F   ......g--agc-cca-c--caacccc--cc----g...a----g-cac---a-cagt-

                              --------- FR3-IMGT ----------------------------------------
B1                            gacacatccaagaaccagttctccctggacctgaactctgtgactcccgaggacacggct
J04097 Homsap IGHV6-1*01 F    --------------------------------c-g-------------------------
Z14223 Homsap IGHV6-1*02 F    --------------------------------c-g-------------------------
Z75355 Homsap IGHV4-4*06 F    -----g-----------------------a-g----g----------cg--c--------c
Z75356 Homsap IGHV4-34*13 F   -----g-----------------------a-g----g----------cg---c-------
Z75359 Homsap IGHV4-59*09 F   -----g-----------------------a-g----g----------cg-t-c-------c

                              ------------>_____ CDR3-IMGT _____
B1                            gtgtattactgtgcaagaggcgttgggtgggagattccagaggcccatggctgtgactac
J04097 Homsap IGHV6-1*01 F    ------------------a
Z14223 Homsap IGHV6-1*02 F    ------------------a
Z75355 Homsap IGHV4-4*06 F    ---------------g----a
Z75356 Homsap IGHV4-34*13 F   ---------------g-----
Z75359 Homsap IGHV4-59*09 F   ---------------g-----
```

Sequence B1

IMGT/V-QUEST provides warnings to alert the users, for the possibility that potential insertions or deletions are present in the sequence

http://www.imgt.org/IMGT_vquest/share/textes/IMGTvquest-warnings.pdf

Warning	Objective
V-GENE and allele: **low V-REGION identity percentage** with the closest V germline	In order to alert users that the sequence has less than 85% of identity with the closest germline V- GENE and allele: this may indicate that the V-GENE and allele name could be not reliable. This could be due to insertion(s) and/or deletion(s).
V-GENE and allele: **2nd-CYS 104 not identified**	In order to alert users that the conserved 2nd-CYS at position 104 according to the unique IMGT numbering is not identified. This could be due to insertion(s) and/or deletion(s).
V-GENE and allele: **different CDR1-IMGT and/or CDR2-IMGT amino acid lengths** compared to closest V germline	In order to alert users that the sequence has different lengths of CDR1 and/or CDR2 compared with theclosest V germline: this may indicate that the V gene and allele name could be not reliable. This could be due to insertion(s) and/or deletion(s).

Sequence B1

- **Back to the submission form**

- **Enable the option at the "advanced parameters" tab**

- **Note that now you can simultaneously analyze a batch of max. 10 sequences**

Advanced parameters

Selection of IMGT reference directory set	F+ORF+ in-frame P	⦿ With all alleles ○ With allele *01 only
Search for insertions and deletions in V-REGION	⦿ Yes (slower, the nb of submitted sequences in a single run is limited to 10)	○ No
Parameters for IMGT/JunctionAnalysis	Nb of accepted D-GENE in IGH JUNCTION (default is 1) default ▼	Nb of accepted mutations: default ▼ in 3'V-REGION default ▼ in D-REGION default ▼ in 5'J-REGION
Parameters for "Detailed view"	Nb of nucleotides to exclude in 5' of the V-REGION for the evaluation of the nb of mutations (in results 9 and 10)	Nb of nucleotides to add (or exclude) in 3' of the V-REGION for the evaluation of the alignment score (in results 1)

Sequence B1

Sequence: 1 B1

Sequence compared with the human IG set from the IMGT reference directory

>B1
```
ccatctccggcgacagtgtctctaccaacagtgctgcttggaactggatcaggcagtccc
catcgagaggccttgagtggctgggaaggacatactacaggAACtcaactggttaaatg
attatgcagtttctgtcaaaggtcgaatagtcatctatgagccctatcaagaaccagt
tctccctggacctgaactctgtgactcccgaggacacggctgtgtattactgtgcaagag
gcgttgggtgggagattccagaggcccatggctgtgactactggggccagggaaccctgg
tcaccgtcttcctcaggta
```

Result summary:	☞ Nucleotide insertions have been detected and automatically removed for this analysis: they are displayed as capital letters in the user submitted sequence above.					
	localization in V-REGION	nb of inserted nt	inserted nt	causing frameshift	from V-REGION codon	from nt position in user submitted sequence
	CDR2-IMGT	3	AAC	no	60	102

	IMGT/V-QUEST results after removal of the insertion(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	Homsap IGHV6-1*01 F, or Homsap IGHV6-1*02 F		score = 1019	identity = 93.91% (216/230 nt) [93.48% (215/230 nt)]
J-GENE and allele	Homsap IGHJ4*02 F		score = 159	identity = 81.25% (39/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD2-15*01 F		D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[2.17.38.11]		[10.9.16]	CARGVGWEIPEAHGCDYW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                             Score      Identity
J04097 Homsap IGHV6-1*01 F   1019       93.91% (216/230 nt)
Z14223 Homsap IGHV6-1*02 F   1019       93.91% (216/230 nt)
M29812 Homsap IGHV4-59*02 F  437        66.97% (146/218 nt)
AB019438 Homsap IGHV4-59*01 F 428       66.51% (145/218 nt)
M95117 Homsap IGHV4-59*04 F  428        66.51% (145/218 nt)
```

3nt insertion detected in VH CDR2 region

displayed in capital letters in the submitted sequence and is **automatically removed** for this analysis

not causing **frameshift** → **Potentially productive** IGH rearranged sequence

% identity to germline **in brackets**, is calculated considering the insertion as **one mutational event**

Sequence B1

The alignment starts from codon VH FR1-24

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

	Score	Identity
J04097 Homsap IGHV6-1*01 F	1019	93.91% (216/230 nt)
Z14223 Homsap IGHV6-1*02 F	1019	93.91% (216/230 nt)
M29812 Homsap IGHV4-59*02 F	437	66.97% (146/218 nt)
AB019438 Homsap IGHV4-59*01 F	428	66.51% (145/218 nt)
M95117 Homsap IGHV4-59*04 F	428	66.51% (145/218 nt)

Alignment with FR-IMGT and CDR-IMGT delimitations

```
                                  <-------------------------------- FR1-IMGT ---------------
                                  ....................................
B1
J04097 Homsap IGHV6-1*01 F        caggtacagctgcagcagtcaggtcca...ggactggtgaagccctcgcagaccctctca
Z14223 Homsap IGHV6-1*02 F        caggtacagctgcagcagtcaggtccg...ggactggtgaagccctcgcagaccctctca
M29812 Homsap IGHV4-59*02 F       caggtgcagctgcaggagtcgggccca...ggactggtgaagccttcggagaccctgtcc
AB019438 Homsap IGHV4-59*01 F     caggtgcagctgcaggagtcgggccca...ggactggtgaagccttcggagaccctgtcc
M95117 Homsap IGHV4-59*04 F       caggtgcagctgcaggagtcgggccca...ggactggtgaagccttcggagaccctgtcc

                                  ------------------->_____ CDR1-IMGT _____<-----
B1                                ..........ccatctccggcgacagtgtctct......accaacagtgctgcttggaac
J04097 Homsap IGHV6-1*01 F        ctcacctgtg-----------g---------......-g--------------
Z14223 Homsap IGHV6-1*02 F        ctcacctgtg-----------g---------......-g--------------
M29812 Homsap IGHV4-59*02 F       ctcacctgca-tg----t--t-g-tcc---.........-gt---tactac----g-
AB019438 Homsap IGHV4-59*01 F     ctcacctgca-tg----t--t-g-tcca--.........-gt---tactac----g-
M95117 Homsap IGHV4-59*04 F       ctcacctgca-tg----t--t-g-tcca--.........-gt---tactac----g-

                                  --------------- FR2-IMGT -------------------->_____ CDR
B1                                tggatcaggcagtccccatcgagaggccttgagtggctgggaaggacatactacaggtcc
J04097 Homsap IGHV6-1*01 F        ------------------------------------------------------------
Z14223 Homsap IGHV6-1*02 F        ------------------------------------------------------------
M29812 Homsap IGHV4-59*02 F       ------c-----c----gg--ag--a--g------a-t--gtat-tc--t-----t...
AB019438 Homsap IGHV4-59*01 F     ------c-----c----gg--ag--a--g------a-t--gtat-tc--t-----t...
M95117 Homsap IGHV4-59*04 F       ------c-----c----gg--ag--a--g------a-t--gtat-tc--t--t--t...

                                  2-IMGT _____<-------------------------------------------
B1                                ...aactggttaaatgattatgcagtttctgtcaaa...ggtcgaatagtcatctatcca
J04097 Homsap IGHV6-1*01 F        ....--g----at--------------a------g--...a---------ac----a-c---
Z14223 Homsap IGHV6-1*02 F        ....--g----at--------------a------g--...a---------ac----a-c---
M29812 Homsap IGHV4-59*02 F       ......g--agc-cca-c--caaccccc--cc----g...a------g-cac---a-cagt-
AB019438 Homsap IGHV4-59*01 F     ......g--agc-cca-c--caaccccc--cc----g...a------g-cac---a-cagt-
M95117 Homsap IGHV4-59*04 F       ......g--agc-cct-c--caacccg--cc----g...a------g-cac---g-cagt-

                                  --------- FR3-IMGT ------------------------------------------
B1                                gacacatccaagaaccagttctccctggacctgaactctgtgactcccgaggacacggct
J04097 Homsap IGHV6-1*01 F        -----------------------------------c-g----------------------
Z14223 Homsap IGHV6-1*02 F        -----------------------------------c-g----------------------
M29812 Homsap IGHV4-59*02 F       -----g---------------------a-g----g---------cg-t-c---------c
AB019438 Homsap IGHV4-59*01 F     -----g---------------------a-g----g---------cg-t-c---------c
M95117 Homsap IGHV4-59*04 F       -----g---------------------a-g----g---------cg---ca---------

                                  ----------->_____ CDR3-IMGT _____
B1                                gtgtattactgtgcaagaggcgttgggtgggagattccagaggcccatggctgtgactac
J04097 Homsap IGHV6-1*01 F        --------------------a
Z14223 Homsap IGHV6-1*02 F        --------------------a
M29812 Homsap IGHV4-59*02 F       --------------g----a
AB019438 Homsap IGHV4-59*01 F     --------------g----a
M95117 Homsap IGHV4-59*04 F       --------------g
```

Sequence B1

Assignment of both *01 and *02 alleles of IGHV6-1

From IMGT Repertoires: Alignments of alleles

http://www.imgt.org/IMGTrepertoire/Proteins/taballeles/human/IGH/IGHV/Hu_IGHVall.html

Alleles *01 and *02 differ only the last position of codon VH FR1-9

Sequence B1

IGHV6-1*01
IGHD2-15*01
IGHJ4*02
identity 93.48%
in-frame
VH CDR3 length 16

→ productive, mutated IGH rearrangement

Practical exercises – Solutions – Batch II

>Sequence B2

gaggtgcacctggtggagtctgggggaggcctggtcaacccggggggggtccctcag
actctcctgtgcagcctctggattcaccttcagtaattttagcatgacctgggtccgcc
aggctccagggcaggggctggagtgggtctcatccatggagtgggtctcatccatta
gtagtagtagtaattacatatactacgcagactcagtgaagggccgattcaccatct
ccagagacaacgccaagaactcactgtatctgcaattgaacagcctgagaggcga
ggacacggctgtttattactgtgcgagaatggtggcttccaagtacccaccatactac
tttgacttctggggcccgggaaccctggtcaccgtctcctcag

Sequence B2

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction) (a)		
V-GENE and allele	Homsap IGHV3-21*02 F	score = 490	identity = 63.54% (183/288 nt)
J-GENE and allele	Homsap IGHJ4*02 F	score = 222	identity = 95.83% (46/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD5-12*01 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.15]	CARMVASKYPPYYFDFW

(a) Low V-REGION identity (63.54%): this may indicate potential nucleotide insertion(s) and/or deletion(s). Try 'Search for insertions and deletions' in 'Advanced parameters' at the bottom of the Search page

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                 Score        Identity
M99658  Homsap IGHV3-21*02 F     490          63.54% (183/288 nt)
AB019439 Homsap IGHV3-21*01 F    481          63.19% (182/288 nt)
AJ879485 Homsap IGHV3-69-1*02 P  475          63.16% (180/285 nt)
KC713940 Homsap IGHV3-11*06 F    472          62.85% (181/288 nt)
HM855323 Homsap IGHV3-21*03 F    472          62.85% (181/288 nt)
```

Sequence B2

Sequence compared with the human IG set from the IMGT reference directory

>B2
gaggtgcacctggtggagtctggggggaggcctggtcaacccggggggggtccctcagactc
tcctgtgcagcctctggattcaccttcagtaattttagcatgacctgggtccgccaggct
ccagggcagggctgGAGTGGGTCTCATCCATGgagtgggtctcatccattagtagtagt
agtaattacatatactacgcagactcagtgaagggccgattcaccatctccagagacaac
gccaagaactcactgtatctgcaattgaacagcctgagaggcgaggacacggctgtttat
tactgtgcgagaatggtggcttccaagtacccaccatactactttgacttctggggcccg
ggaaccctggtcaccgtctcctcag

Result summary:	Nucleotide insertions have been detected and automatically removed for this analysis: they are displayed as capital letters in the user submitted sequence above.					
	localization in V-REGION	nb of inserted nt	inserted nt	causing frameshift	from V-REGION codon	from nt position in user submitted sequence
	FR2-IMGT	18	GAGTGGGTCTCATCCATG	no	51	136

IMGT/V-QUEST results after removal of the insertion(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	Homsap IGHV3-21*01 F, or Homsap IGHV3-21*02 F	score = 1318	identity = 95.49% (275/288 nt) [95.14% (274/288 nt)]
J-GENE and allele	Homsap IGHJ4*02 F	score = 222	identity = 95.83% (46/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD5-12*01 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.15]	CARMVASKYPPYYFDFW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                  Score       Identity
AB019439 Homsap IGHV3-21*01 F     1318        95.49% (275/288 nt)
M99658 Homsap IGHV3-21*02 F       1318        95.49% (275/288 nt)
HM855323 Homsap IGHV3-21*03 F     1309        95.14% (274/288 nt)
HM855688 Homsap IGHV3-21*04 F     1309        95.14% (274/288 nt)
HM855336 Homsap IGHV3-48*04 F     1246        92.71% (267/288 nt)
```

Sequence B2

Assignment of both *01 and *02 alleles of IGHV3-21

From IMGT Repertoires: Alignments of alleles

http://www.imgt.org/IMGTrepertoire/Proteins/taballeles/human/IGH/IGHV/Hu_IGHVall.html

Alleles *01 and *02 differ only the last position of codon VH FR1-3

Sequence B2

Assignment of both *01 and *02 alleles of IGHV3-21

The position where the alleles differ is mutated in this rearrangement

6. V-REGION alignment according to the IMGT unique numbering
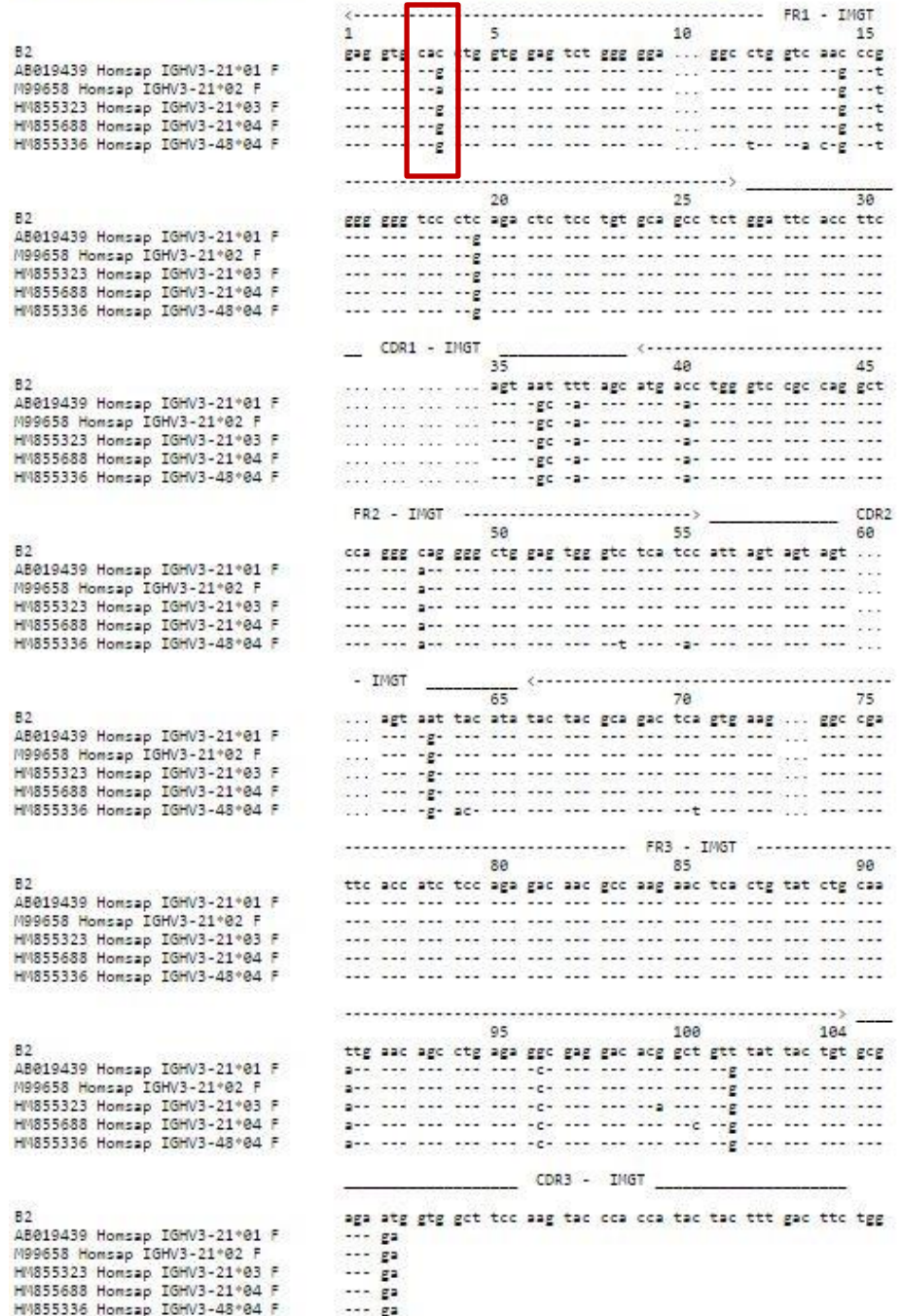
Sequence B2

Analysis of the JUNCTION

D-REGION is in reading frame 1.

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N1	D-REGION	N2	P	5'J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
B2	Homsap IGHV3-21*01	tgtgcgaga..	......atggtggct........	tccaagtacccacca	t	actactttgactтctgg	Homsap IGHJ4*02	Homsap IGHD5-12*01	0	1	1	8/15

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	110	111	111.1	112.1	112	113	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	A	R	M	V	A	S	K	Y	P	P	Y	Y	F	D	F	W					
B2	tgt	gcg	aga	atg	gtg	gct	tcc	aag	tac	cca	cca	tac	tac	ttt	gac	ttc	tgg	+	15	2,144.51	8.19	CARMVASKYPPYYFDFW

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

Sequence B2

IGHV3-21*01
IGHD5-12*01
IGHJ4*02
identity 95.14%
in-frame
VH CDR3 length 15

→ productive, mutated IGH rearrangement

Practical exercises – Solutions – Batch II

>Sequence B3

gaggtgcacctggtggagtctgggggaggcctggtcaacccggggggggtccctcag
actctcctgtgcagcctctggattcaccttcagtaattttagcatgacctgggtccgcc
aggctccagggcaggggctggagtgggtctcatccatggagtgggtctcatccatta
gtagtagtagtaattacatatactacgcagactcagtgaagggccgattcaccatct
ccagagacaacgccaagaactcactgtatctgcaattgaacagcctgagaggcga
ggacacggctgtttattactgtgcgagaatggtggcttccaagtacccaccatactac
tttgacttctggggcccgggaaccctggtcaccgtctcctcag

Sequence B3

Sequence: 1 B3

Sequence compared with the human IG set from the IMGT reference directory

>B3
gtgggtccttcagtggttactattggacttggatccgtcagcccccagggaaggggctgg
agtggattggggaaatcaatcatagtggactagggaaggggctggagtggattgcgaaaa
tcagtcatagtgaaactgcgaactacaacccgtccttcaagagtcgagtcaccatatcac
tagacacgtccaagaaccagttctccctgaacttgagatctgtgaccgccgggggatacgg
ctgtatattactgtgcgagaggccttcaagttggtgggacctactctcttgattattggg
gccagggaaccctggtcatcgtctcctcag

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction) (a)		
V-GENE and allele	Homsap IGHV4-34*13 F	score = 547	identity = 73.24% (156/213 nt)
J-GENE and allele	Homsap IGHJ4*02 F	score = 186	identity = 87.50% (42/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD1-26*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[14.17.38.11]	[8.7.14]	CARGLQVGGTYSLDYW

(a) Low V-REGION identity (73.24%): this may indicate potential nucleotide insertion(s) and/or deletion(s). Try 'Search for insertions and deletions' in 'Advanced parameters' at the bottom of the Search page

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                    Score       Identity
Z75356 Homsap IGHV4-34*13 F         547         73.24% (156/213 nt)
Z75351 Homsap IGHV4-30-2*04 F       502         70.89% (151/213 nt)
Z75347 Homsap IGHV4-61*06 ORF       484         69.95% (149/213 nt)
Z75355 Homsap IGHV4-4*06 F          478         69.01% (147/213 nt)
Z75353 Homsap IGHV4-30-4*05 F       475         69.16% (148/214 nt)
```

Sequence B3

Sequence: 1 B3

Sequence compared with the human IG set from the IMGT reference directory

>B3

gtgggtccttcagtggttactattggacttggatccgtcagccccccaggggaaggggctgg
agtggattggggaaatcaatcatagtggaCTAGGGAAGGGGCTGGAGTGGATTGCGAAAA
TCAGTCATAGTGAAactgcgaactacaacccgtccttcaagagtcgagtcaccatatcac
tagacacgtccaagaaccagttctccctgaacttgagatctgtgaccgccgggggatacgg
ctgtatattactgtgcgagaggccttcaagttggtgggacctactctcttgattattggg
gccaggggaaccctggtcatcgtctcctcag

Result summary:	Nucleotide insertions have been detected and automatically removed for this analysis: they are displayed as capital letters in the user submitted sequence above.					
	localization in V-REGION	nb of inserted nt	inserted nt	causing frameshift	from V-REGION codon	from nt position in user submitted sequence
	CDR2-IMGT	45	CTAGGGAAGGGGCTGGAGTGGATTGCGAAAATCAGTCATAGTGAA	no	84	90

IMGT/V-QUEST results after removal of the insertion(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	Homsap IGHV4-34*01 F, or Homsap IGHV4-34*02 F	score = 896	identity = 92.34% (193/209 nt) [91.87% (192/209 nt)]
J-GENE and allele	Homsap IGHJ4*02 F	score = 186	identity = 87.50% (42/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD1-26*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[X.17.38.11]	[7.7.14]	CARGLQVGGTYSLDYW

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

	Score	Identity
AB019439 Homsap IGHV4-34*01 F	896	92.34% (193/209 nt)
M99684 Homsap IGHV4-34*02 F	896	92.34% (193/209 nt)
X92255 Homsap IGHV4-34*03 F	887	91.83% (191/208 nt)
M95113 Homsap IGHV4-34*08 F	887	91.87% (192/209 nt)
X56591 Homsap IGHV4-34*12 F	887	91.87% (192/209 nt)

This rearrangement is productive.

The insertion does not cause frameshift, but its length (45nt) may render it non functional

Sequence B3

IGHV4-34*01
IGHD1-26*01
IGHJ4*02
identity 91.87%
in-frame
VH CDR3 length 14

→ productive, mutated IGH rearrangement

Practical exercises – Solutions – Batch II

>**Sequence B4**

gaggaggaggtgttggagtctgggggaggcttggtacagcctggggggtccctgag
actctcctgtgtagcctctggattcacctttagtaattatgtcatgagttgggtccgcc
aggctccagggaggggactggagtgggtctccgttattactgatagtggtggtcata
catactatgcagactccgtgaagggccggttcaccatctccagagacaattccaaga
acaccctatatcttcaaatgaacagcctgagagtcgaggacacggccagatattact
gtgcgagagactggggggatattgtggggggaggcaccctggtcaccgtctccaca
t

Sequence B4

Sequence: 1 B4

Sequence compared with the human IG set from the IMGT reference directory

>B4
gaggaggaggtgttggagtctgggggaggcttggtacagcctggggggtccctgagactc
tcctgtgtagcctctggattcacctttagtaattatgtcatgagttgggtccgccaggct
ccaggaaggggactggagtgggtctccgttattactgatagtggtggtcatacatactat
gcagactccgtgaagggccggttcaccatctccagagacaattccaagaacaccctatat
cttcaaatgaacagcctgagagtcgaggacacggccagatattactgtgcgagagactgg
ggggatattgtggggggaggcaccctggtcaccgtctccacat

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV3-23*01 F, or Homsap IGHV3-23D*01 F	score = 1210	identity = 91.32% (263/288 nt)
J-GENE and allele	Homsap IGHJ5*01 F	score = 100	identity = 68.09% (32/47 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD7-27*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.7]	CARDWGDIV (TRP 118 not identified)

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                    Score    Identity
M99660 Homsap IGHV3-23*01 F         1210     91.32% (263/288 nt)
AC244492 Homsap IGHV3-23D*01 F      1210     91.32% (263/288 nt)
AJ879486 Homsap IGHV3-23*04 F       1201     90.97% (262/288 nt)
AC245166 Homsap IGHV3-23D*02 F      1201     90.97% (262/288 nt)
M35415 Homsap IGHV3-23*02 F         1192     90.62% (261/288 nt)
```

The 2nd VH CDR3 anchor is not identified and the VH CDR3 length is unusually short

Sequence B4

4. Results of IMGT/JunctionAnalysis

Maximum number of accepted mutations in: 3'V-REGION = 2, D-REGION = 4, 5'J-REGION = 2
Maximum number of accepted D-GENE = 1

Analysis of the JUNCTION

D-REGION is in reading frame 3.

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N1	D-REGION	N2	5'J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
B4	Homsap IGHV3-23*01	tgtgcgagaga		....ctgggg.	ggatattgt	....................g	Homsap IGHJ5*01	Homsap IGHD7-27*01	1	0	0	3/9

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	A	R	D	W	G	D	I	V					
B4	tgt	gcg	aga	gac	tgg	ggg	gat	att	gtg	+	7	1,034.16	4.44	CARDWGDIV

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

Non identification of W118 is owed to extreme 5'-3' exonuclease activity

Sequence B4

The amplification reaction should be repeated using different sets of primers in an attempt to amplify a second rearrangement.

However, this mutated codon 118 is followed by G-Q-G motif (IMGT positions 119-121) in the FR4 encoded by the IGHJ5 gene,

The rearrangement is assumed to be productive and can be used for interpretation of the mutational status.

7. V-REGION translation

Sequence B4

IGHV3-23*01 or IGHV3-23D*01
IGHD7-27*01
IGHJ5*01
identity 91.32%
in-frame
VH CDR3 length 7

**→ productive, mutated IGH
rearrangement**

Practical exercises – Solutions – Batch II

>Sequence B5

gaggtgcagctggtggagtctgggggaggcttggtccagccggggggggtccctgag
actctcctgtgcagcctctggattcacctttaataggtattggatgagttgggtccgcc
aggctccagggaaggggctggagtgggtggccaacataaaggaagatggaagtg
agaaatactatgtggactctgtgaagggccgattcaccatctccagagacaacgac
aagaactcactgtatctgcaaatgaacagtctgagagtcgaagacacggccgtgta
ctactgtgcgagagtccttttctttggagaggttatttgggggggtcactggttcgacc
cccggggccagggaaccct

Sequence B5

Sequence: 1 B5

Sequence compared with the human IG set from the IMGT reference directory

>B5

gaggtgcagctggtggagtctgggggaggcttggtccagccggggggggtccctgagactc
tcctgtgcagcctctggattcacctttaataggtattggatgagttgggtccgccaggct
ccagggaaggggctggagtgggtggccaacataaaggaagatggaagtgagaaatactat
gtggactctgtgaagggccgattcaccatctccagagacaacgacaagaactcactgtat
ctgcaaatgaacagtctgagagtcgaagacacggccgtgtactactgtgcgagagtcctt
ttctttggagaggttatttgggggggtcactggttcgacccccggggccaggggaaccct

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV3-7*03 F	score = 1345	identity = 96.53% (278/288 nt)
J-GENE and allele	Homsap IGHJ5*02 F	score = 134	identity = 88.24% (30/34 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD3-3*01 F	D-REGION is in reading frame 3	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.5]	[8.8.18]	CARVLFFGEVIWGGHWFDPR (TRP 118 not identified)

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

		Score	Identity
HM855666	Homsap IGHV3-7*03 F	1345	96.53% (278/288 nt)
M99649	Homsap IGHV3-7*01 F	1336	96.18% (277/288 nt)
X92288	Homsap IGHV3-7*02 F	1327	95.83% (276/288 nt)
M99678	Homsap IGHV3-52*01 P	1129	88.19% (254/288 nt)
Z17388	Homsap IGHV3-52*02 P	1120	87.85% (253/288 nt)

The 2nd VH CDR3 anchor is not identified

Sequence B5

Analysis of the JUNCTION

D-REGION is in reading frame 3.

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N1	D-REGION	N2	5'J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
B5	Homsap IGHV3-7*03	tgtgcgagag.	tccttttc	...........tttggagaggttatt......	tgggggggtc	...actggttcgacccccgg	Homsap IGHJ5*02	Homsap IGHD3-3*01	0	1	1	11/18

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	110	111	111.1	111.2	112.3	112.2	112.1	112	113	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	A	R	V	L	F	F	G	E	V	I	W	G	G	H	W	F	D	P	R					
B5	tgt	gcg	aga	gtc	ctt	ttc	ttt	gga	gag	gtt	att	tgg	ggg	ggt	cac	tgg	ttc	gac	ccc	cgg	+	18	2,392.77	7.18	CARVLFFGEVIWGGHWFDPR

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

Non identification of W118 is owed to SHM activity

Sequence B5

The VH FR4 GXG motif is present in the rearrangement

The rearrangement is assumed to be productive and can be used for interpretation of the mutational status.

Sequence B5

IGHV3-7*03
IGHD3-3*01
IGHJ5*02
identity 96.53%
in-frame
VH CDR3 length 18

→ productive, mutated IGH rearrangement

Practical exercises – Solutions – Batch II

>Sequence B6

ctgtgttgacgcagccgccttctgcatcagcctccctgggagcctcggtcacactcac
ctgcaccctgagcagcgactaaagtaattataaagtggactggtaccaagagagac
cagggcagggcccccagtttgtgatgcgagtgggcactggtgggattgtgggatcc
aaggggatggcatccctgatcgcttctcagtcttggactcaggcctgagtcggttgc
tgaccatcaagaacatccaagaagaggatgagagtgactaccactgtggggcaga
ccatggcgatgggaacaacttcgtcccctgggtgttaggcggagggaccaagctga
ccgtcctag

Sequence B6

Sequence B6

4. Results of IMGT/JunctionAnalysis

Maximum number of accepted mutations in: 3'V-REGION = 7, 5'J-REGION = 7
Maximum number of accepted D-GENE = 0

Analysis of the JUNCTION

Click on mutated (underlined) nucleotide to see the original one:

Input	V name	3'V-REGION	N	5'J-REGION	J name	Vmut	Jmut	Ngc
B6	Homsap IGLV9-49*01	tgtggggcagaccatggcgatgggaacaacttcgt......	cccc	.tgggtgtta	Homsap IGLJ3*02	3	1	4/4

Translation of the JUNCTION

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	110	111	112.1	112	113	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass	pI	PhysicoChemical Descriptor (by BRFAA)
	C	G	A	D	H	G	D	G	N	N	F	V	P	W	V	L					
B6	tgt	ggg	gca	gac	cat	ggc	gat	ggg	aac	aac	ttc	gtc	ccc	tgg	gtg	tta	+	14	1,700.85	4.43	CGADHGDGNNFVPWVL

Be aware that some allele reference sequences may be incomplete or from cDNAs. In those cases, IMGT/JunctionAnalysis uses automatically the allele *01 for the analysis of the JUNCTION.

Non identification of F118 is owed to SHM activity

Sequence B6

The rearrangement is rendered unproductive due to the presence of a stop codon in position VL CDR1-29

Sequence B6

IGLV9-49*01
IGLJ3*02
identity 94.74%
in-frame
VH CDR3 length 18

→ unproductive, mutated IGL rearranged sequence

Practical exercises – Solutions – Batch II

>Sequence B7
cccagactccactctccttgtctatcacccctggagagcaggcctccatgtcctgcag
gtctagtcagagcctcctgcatagtgatggatacacctatttgtattggtttctgcaga
aagccaggccagtctccacagctcctgatctatgaagtttccaaccggttctctggag
tgccagataggttcagtggcagcgggtcagggacagatttcacactgaaaatcagc
cgggtggaggctgaggatgttggagtttattactgcatgcaagatgcacaagatcct
cggctttcggcggagggaccaagctggagatcaaacgt

Sequence B7

Sequence compared with the human IG set from the IMGT reference directory

>B7
cccagactccactctccttgtctatcacccctggagagcaggcctccatgtcctgcaggt
ctagtcagagcctcctgcatagtgatggatacacctatttgtattggtttctgcagaaag
ccaggccagtctccacagctcctgatctatgaagtttccaaccggttctctggagtgcca
gataggttcagtggcagcgggtcagggacagatttcacactgaaaatcagccgggtggag
gctgaggatgttggagtttattactgcatgcaagatgcacaagatcctcggctttcggcg
gagggaccaagctggagatcaaacgt

Result summary:	Unproductive IGK rearranged sequence (out-of-frame junction) (a)		
V-GENE and allele	Homsap IGKV2-29*03 F, or Homsap IGKV2D-29*02 F	score = 523	identity = 64.89% (183/282 nt)
J-GENE and allele	Homsap IGKJ4*01 F (b)	score = 148	identity = 91.43% (32/35 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[22.17.36.10]	[11.3.X]	CMQDAQDPR#F

(a) Low V-REGION identity (64.89%): this may indicate potential nucleotide insertion(s) and/or deletion(s). Try 'Search for insertions and deletions' in 'Advanced parameters' at the bottom of the Search page

. Other possibilities: Homsap_IGKJ2*01 and Homsap_IGKJ2*02 (highest number of consecutive identical nucleotides)

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon plus 15 nt of the CDR3-IMGT)

```
                                Score      Identity
AJ783437 Homsap IGKV2-29*03 F    523       64.89% (183/282 nt)
U41644 Homsap IGKV2D-29*02 F     523       64.89% (183/282 nt)
X63396 Homsap IGKV2-29*01 P      514       64.54% (182/282 nt)
U41645 Homsap IGKV2-29*02 F      514       64.54% (182/282 nt)
M31952 Homsap IGKV2D-29*01 F     514       64.54% (182/282 nt)
```

Sequence B7

Sequence: 1 B7

Sequence compared with the human IG set from the IMGT reference directory

>B7
cccagactccactctccttgtctatcacccctggagagcaggcctccatgtcctgcaggt
ctagtcagagcctcctgcatagtgatggatacacctatttgtattggtttctgcagaaag
ccaggccagtctccacAgctcctgatctatgaagtttccaaccggttctctggagtgcca
gataggttcagtggcagcgggtcagggacagatttcacactgaaaatcagccgggtggag
gctgaggatgttggagtttattactgcatgcaagatgcacaagatcctcggctttcggcg
gagggaccaagctggagatcaaacgt

	☞ Nucleotide insertions have been detected and automatically removed for this analysis: they are displayed as capital letters in the user submitted sequence above.					
Result summary:	localization in V-REGION	nb of inserted nt	inserted nt	causing frameshift	from V-REGION codon	from nt position in user submitted sequence
	FR2-IMGT	1	A	yes	51	137

IMGT/V-QUEST results after removal of the insertion(s): Unproductive IGK rearranged sequence (out-of-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	Homsap IGKV2D-28*03 F	score = 1391	identity = 99.64% (280/281 nt) [99.29% (279/281 nt)]
J-GENE and allele	Homsap IGKJ4*01 F (a)	score = 148	identity = 91.43% (32/35 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[21.17.36.10]	[11.3.X]	CMQDAQDPR#F

. Other possibilities: Homsap_IGKJ2*01 and Homsap_IGKJ2*02 (highest number of consecutive identical nucleotides)

Repeat/check sequencing reaction/output to decide whether the insertion is true

If this is the single rearrangement amplified repeat the amplification with different sets of primers and/or try to amplify a lamda chain rearrangement in order to:
> (i) confirm the presence of the unproductive rearrangement
> (ii) look for the presence of a second productive rearrangement

Sequence B7

IGKV2D-26*03
IGKJ4*01
identity 99.29%
out-of-frame
VH CDR3 length X

→ unproductive, unmutated IGK rearranged sequence

Practical exercises – Solutions – Batch II

>Sequence B8

cagctgcagctgcaggagtcgggcccaggactggcggagaccctgtccctcacctg
cactgtctctggtggctccatcagctctatcgattactactggggctggatacgccag
cccccagggaaggggctggagtggattgggagtgtctattatagtgggagcaccta
ctacaatccgtccctcaagagtcgagtcaccatgtccgtcgacacgtccaagaacca
gttctccctgaagttgagctctgtgaccgccgcagacacggctgtgtattactgtgtg
agacttccgggtatttcagtgccacaatactacactgtggacgtctggggccaggg
accacggtcaccgtctcctca

Sequence B8

The rearrangement shows low V-gene identity but doesn't activate the IMGT warning

Sequence B8

Poor quality alignment at the beginning of VH FR1 region resulting in **mutational imbalance** compared to the remaining V-region

Enable the "search for indels" option

Alignment with FR-IMGT and CDR-IMGT delimitations

```
                                <------------------------------------- FR1-IMGT ----------------
B8                              ........ cagctgcagctgcaggag...tcgggcccaggactggcggagacctgtcc
AB019439 Homsap IGHV4-39*01 F   cagctgcag-t--a-g--tc-ggccca...ggactggtgaag-ctt------------
X05715 Homsap IGHV4-39*02 F     cagctgcag-t--a-g--tc-ggccca...ggactggtgaag-ctt------------
X92259 Homsap IGHV4-39*03 F     cagctgcag-t--a-g--tc-ggccca...ggactggtgaag-ctt------------
M95116 Homsap IGHV4-39*05 F     cagctgcag-t--a-g--tc-ggccca...ggactggtgaag-ctt---------C----
AM940222 Homsap IGHV4-39*07 F   cagctgcag-t--a-g--tc-ggccca...ggactggtgaag-ctt------------

                                ----------------->_____ CDR1-IMGT _____<-----
B8                              ctcacctgcactgtctctggtggctccatcagc......tctatcgattactactggggc
AB019439 Homsap IGHV4-39*01 F   ------------------------------------......ag--gtag-------------
X05715 Homsap IGHV4-39*02 F     ------------------------------------......ag--gtag-------------
X92259 Homsap IGHV4-39*03 F     ------------------------------------......ag--gtag-------------
M95116 Homsap IGHV4-39*05 F     ------------------------------------......ag--gtag-------------
AM940222 Homsap IGHV4-39*07 F   ------------------------------------......ag--gtag-------------

                                --------------- FR2-IMGT -------------------->_____ CDR
B8                              tggatacgccagcccccaggggaagggggctggagtggattgggagtgtctattatagt...
AB019439 Homsap IGHV4-39*01 F   -----C--------------------------------------a---...
X05715 Homsap IGHV4-39*02 F     -----C--------------------------------------a---...
X92259 Homsap IGHV4-39*03 F     -----C--------------------------------------a---...
M95116 Homsap IGHV4-39*05 F     -----C--------------------------------------a---...
AM940222 Homsap IGHV4-39*07 F   -----C--------------------------------------a---...

                                2-IMGT _____<------------------------------------
B8                              ......gggagcacctactacaatccgtccctcaag...agtcgagtcaccatgtccgtc
AB019439 Homsap IGHV4-39*01 F   ......-------------------C-----------,,,-------------a-----a
X05715 Homsap IGHV4-39*02 F     ......-------------------C-----------,,,-------------a-----a
X92259 Homsap IGHV4-39*03 F     ......-------------------C-----------,,,-------------a-----a
M95116 Homsap IGHV4-39*05 F     ......-------------------C-----------,,,-------------a-----a
AM940222 Homsap IGHV4-39*07 F   ......-------------------C-----------,,,-------------a--a--a

                                --------- FR3-IMGT -----------------------------------------
B8                              gacacgtccaagaaccagttctccctgaagttgagctctgtgaccgccgcagacacggct
AB019439 Homsap IGHV4-39*01 F   --------------------------------C---------------------------
X05715 Homsap IGHV4-39*02 F     ------------------C-------------C---------------------------
X92259 Homsap IGHV4-39*03 F     --------------------------------C-------------------------c
M95116 Homsap IGHV4-39*05 F     --------------------------------C---------------------------
AM940222 Homsap IGHV4-39*07 F   --------------------------------C-----------------g--------c

                                ----------->_____ CDR3-IMGT _____
B8                              gtgtattactgtgtgagacttccggggtatttcagtgccacaatactacactgtggacgtc
AB019439 Homsap IGHV4-39*01 F   --------------C-----a
X05715 Homsap IGHV4-39*02 F     --------------C----ga
X92259 Homsap IGHV4-39*03 F     -----------
M95116 Homsap IGHV4-39*05 F     --------------C-
AM940222 Homsap IGHV4-39*07 F   --------------C----ga
```

Sequence B8

Sequence: 1 B8

Sequence compared with the human IG set from the IMGT reference directory

>B8
cagctgcagctgcaggagtcgggcccaggactggcggagaccctgtccctcacctgcact
gtctctggtggctccatcagctctatcgattactactggggctggatacgccagcccca
gggaaggggctggagtggattgggagtgtctattatagtgggagcacctactacaatccg
tccctcaagagtcgagtcaccatgtccgtcgacacgtccaagaaccagttctccctgaag
ttgagctctgtgaccgccgcagacacggctgtgtattactgtgtgagacttccgggtatt
tcagtgccacaatactacactgtggacgtctggggccaagggaccacggtcaccgtctcc
tca

<table>
<tr><td rowspan="3">Result summary:</td><td colspan="5">👈☞ Nucleotide deletions have been detected (shown by dots in the alignments):</td></tr>
<tr><td>localization</td><td>nb of deleted nt</td><td>causing frame shift</td><td>from V-REGION codon</td><td>from nt position in user submitted sequence</td></tr>
<tr><td>FR1-IMGT</td><td>9</td><td>no</td><td>13</td><td>34</td></tr>
</table>

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	Homsap IGHV4-39*01 F	score = 1288	identity = 95.39% (269/282 nt) [95.04% (268/282 nt)]
J-GENE and allele	Homsap IGHJ6*02 F (a)	score = 193	identity = 79.03% (49/62 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD6-19*01 F	D-REGION is in reading frame 2	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[22.17.38.11]	[10.7.16]	CVRLPGISVPQYYTVDVW

(a) Other possibilities: Homsap_IGHJ3*01 (shorter alignment but highest percentage of identity)

---

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                                Score        Identity
AB019439 Homsap IGHV4-39*01 F   1288         95.39% (269/282 nt)
X05715 Homsap IGHV4-39*02 F     1279         95.04% (268/282 nt)
X92259 Homsap IGHV4-39*03 F     1279         95.02% (267/281 nt)
M95116 Homsap IGHV4-39*05 F     1279         95.04% (268/282 nt)
AM940222 Homsap IGHV4-39*07 F   1261         94.33% (266/282 nt)
```

Sequence B8

Gene assignment is not altered

The V-gene identity score and the alignment is improved

The 9nt deletion is shown in the alignment

Alignment with FR-IMGT and CDR-IMGT delimitations

```
                                     <------------------------------- FR1-IMGT ------------------
B8                                   cagctgcagctgcaggagtcgggccca...ggactg.........gcggagacctgtcc
AB019439 Homsap IGHV4-39*01 F        ----------------------------,,.---------gtgaagcctt----------------
X05715 Homsap IGHV4-39*02 F          ----------------------------,,.---------gtgaagcctt----------------
X92259 Homsap IGHV4-39*03 F          ----------------------------,,.---------gtgaagcctt----------------
M95116 Homsap IGHV4-39*05 F          ----------------------------,,.---------gtgaagcctt---------c----
AM940222 Homsap IGHV4-39*07 F        ----------------------------,,.---------gtgaagcctt----------------
```

```
                                     ---------------->_____ CDR1-IMGT _____<-----
B8                                   ctcacctgcactgtctctggtggctccatcagc......tctatcgattactactggggc
AB019439 Homsap IGHV4-39*01 F        ------------------------------......ag--gtag-------------
X05715 Homsap IGHV4-39*02 F          ------------------------------......ag--gtag-------------
X92259 Homsap IGHV4-39*03 F          ------------------------------......ag--gtag-------------
M95116 Homsap IGHV4-39*05 F          ------------------------------......ag--gtag-------------
AM940222 Homsap IGHV4-39*07 F        ------------------------------......ag--gtag-------------
```

```
                                     --------------- FR2-IMGT --------------------->_____ CDR
B8                                   tggatacgccagcccccagggaaggggctggagtggattgggagtgtctattatagt...
AB019439 Homsap IGHV4-39*01 F        -----c------------------------------------a---------...
X05715 Homsap IGHV4-39*02 F          -----c------------------------------------a---------...
X92259 Homsap IGHV4-39*03 F          -----c------------------------------------a---------...
M95116 Homsap IGHV4-39*05 F          -----c------------------------------------a---------...
AM940222 Homsap IGHV4-39*07 F        -----c------------------------------------a---------...
```

```
                                     2-IMGT _____<------------------------------------------
B8                                   ......gggagcacctactacaatccgtccctcaag...agtcgagtcaccatgtccgtc
AB019439 Homsap IGHV4-39*01 F        ......,---------------------c-------,,----------a-----a
X05715 Homsap IGHV4-39*02 F          ......,---------------------c-------,,----------a-----a
X92259 Homsap IGHV4-39*03 F          ......,---------------------c-------,,----------a-----a
M95116 Homsap IGHV4-39*05 F          ......,---------------------c-------,,----------a-----a
AM940222 Homsap IGHV4-39*07 F        ......,---------------------c-------,,----------a--a--a
```

```
                                     --------- FR3-IMGT ----------------------------------------
B8                                   gacacgtccaagaaccagttctccctgaagttgagctctgtgaccgccgcagacacggct
AB019439 Homsap IGHV4-39*01 F        ----------------------------------c-------------------
X05715 Homsap IGHV4-39*02 F          ----------------c-----------------c-------------------
X92259 Homsap IGHV4-39*03 F          ----------------------------------c-------------------c
M95116 Homsap IGHV4-39*05 F          ----------------------------------c-------------------
AM940222 Homsap IGHV4-39*07 F        ----------------------------------c----------g--------c
```

```
                                     ----------->_____ CDR3-IMGT _____
B8                                   gtgtattactgtgtgagacttccgggtatttcagtgccacaatactacactgtggacgtc
AB019439 Homsap IGHV4-39*01 F        -------------c-----a
X05715 Homsap IGHV4-39*02 F          -------------c----ga
X92259 Homsap IGHV4-39*03 F          -----------
M95116 Homsap IGHV4-39*05 F          -------------c-
AM940222 Homsap IGHV4-39*07 F        -------------c----ga
```

Sequence B8

IGHV4-39*01
IGHD6-19*01
IGHJ6*02
identity 95.04%
in-frame
VH CDR3 length 16

→ productive, mutated IGH rearrangement