IMGT databases and tools for the analysis of the B cell repertoire in normal and pathological situations

Marie-Paule Lefranc

Laboratoire d’ImmunoGénétique Moléculaire
Université Montpellier, UPR CNRS 1142, IGH
Institut Universitaire de France

Colloque d’Oléron, Lymphocytes B normaux et tumoraux,
May 12-15, 2005, Ile d’Oléron, France
When and why was IMGT created?
1. IMGT was created in June 1989, following the Human Genome Meeting HGM, at New Haven, USA.
   * Thomas Shows and Ken Kidd, meeting organizers
   * Lap Chee Tsui, chromosome 7 responsable
   * Paul Boucheix (CD nomenclature), Roland Berger (translocations).

At New haven, entry of the 16 human T cell receptor gamma genes (TRG) in the genome database, Genome Databank (GDB).

2. IMGT was created to answer the need to manage the T cell receptor (TR) and immunoglobulin (IG) genes and to enter them in the generalist databases.
Why was IMGT necessary?

1. Unique structure of the immunoglobulin (IG) and T cell receptor (TR) genes.
Genomic DNA in germline configuration

V-GENE

>XM2106.0|HSV12|Homo sapiens VI-2 gene for immunoglobulin heavy chain

tgagagctcc gttcctcacc **atggactgga** cctggaggat cctctttcttg gtggcagcag 60
ccacaggca gaggctccct agtcccgagtg atgagaaga gaattgactcc agtccaggga 120
gatctcactc acttctgtgt tctctccaga **ggaggccccact** cccaggggca gctgggtgcag 180
tctggggcttg aggtgaagga gcctgggccc tcagtgaaag tgtctctgca ggcctctttga 240
tagacctca cccgctacta tatgcactgg gtcgagccagag cccctggaag aagggctttga 300
tgatggtgggt gaatactgcc taacagtgtg gcacacaact agcagcagaa gtttcagggcg 360
agggctggtc cagcaggaga cagctgtccatc agcagcagct acatggagct gacaggcttg 420
agatctcgacg cacgagggcgt gtattactggt gcgagagaca cagtgtgaaag acccacatcc 480
tgagggtgtgc agaaaccdd a gggaggaggg cag

http://imgt.cines.fr
Genomic DNA in germline configuration

**D-GENE**

> J00256 | IGH7-27*01 | Homo sapiens D-GENE

ccagccgcag gttttttgtgc ttagctgaga acsactgtgc taactgggga cacagtgatt
ggcagctcta caaaaaaccat gtcctcccccgg g

**J-GENE**

> J00256 | IGHJ1*01 | Homo sapiens J-GENE

accccggggtc gtgggtttct gtgccccctgg ctcagggtctg actcaccgtg gctgaatact
tccagcacatg gggccagggc accctgttct ca cctctcctc aggtgagtct gctgtactgg
ggatagcggg gagccatgtg tactgggcca agcaaggct ttggcttcag

http://imgt.cines.fr
2. Huge diversity of the immunoglobulin and T cell receptor repertoires.
What was the first goal?

CONTENT

1. Immunoglobulins AND T cell receptors
2. Human AND other vertebrate species

DATA

1. Genes AND alleles
2. Sequences (germline AND rearranged)
3. Polymorphisms: genomic AND somatic hypermutations
4. 3D structures
Spacefill 3D representation of an IgG

- **VH**: V-DOMAIN of the immunoglobulin heavy chain
- **VL**: V-DOMAIN of the immunoglobulin light chain
- **CH1, CH2, CH3**: C-DOMAIN of the immunoglobulin heavy chain
- **CL**: C-DOMAIN of the immunoglobulin light chain

VH corresponds to the V-D-J-REGION (in green (V), orange (DJ)) of the heavy chain.

VL corresponds to the V-J-REGION (in green (V) and yellow (DJ)) of the light chain.

(The Immunoglobulin FactsBook, 2001)
Immunoglobulin (IG)  

T cell receptor (TR)

Contribution of the 2 V-DOMAINs to the antigen binding site

Alpha - Beta
Gamma - Delta

Membrane IgM

T cell receptor
Immunoglobulin IgG

Immunoglobulin (IG) synthesis

- Genomic DNA (IGH Locus 14q32)
  - V\(\rightarrow\)D\(\rightarrow\)J\(\rightarrow\)C
  - 5' 3'

- Arranged DNA

- mRNA

- 2 x 10^{12} different IG per individual

Immunoglobulin (IG) synthesis

150 FUNCTIONAL IG GENES

HEAVY CHAIN

V  5'
D  39-46
J  x 23
C  x 6

5' 3'

LIGHT CHAIN

V  5'
J  34-37
C  x 5

Kappa

3' 30-33
Lambda

x 4-5

6300 POTENTIAL RECOMBINATIONS

N-DIVERSITY
SOMATIC MUTATIONS
x 1000

ABOUT 6.3 x 10^6 POSSIBILITIES

2 x 10^{12} DIFFERENT ANTIBODIES

185 +165 POTENTIAL RECOMBINATIONS

ABOUT 3.5 x 10^5 POSSIBILITIES

IMGT® Web resources

**IMGT Repertoire** 8,000 pages HTML
- Chromosomal localizations
- Locus representations,
- Genes tables
- Potential repertoire
- Alignments of alleles
- Protein displays
- IMGT Colliers de Perles...

**IMGT Index**

**IMGT Bloc-notes**
- Interesting links, PubMed,
- Meeting announcements
- IMGT Immunoinformatics page,...

**IMGT Education**
- Tutorials, IMGT Lexique, ...
### Human IGHV genes on chromosome 14 at 14q32.33

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Functional</th>
<th>ORF</th>
<th>Pseudogene</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGHV1</td>
<td>9</td>
<td>-</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>IGHV2</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IGHV3</td>
<td>18-20**(1)**</td>
<td>3</td>
<td>25**(1)**</td>
<td>47-49**</td>
</tr>
<tr>
<td>IGHV4</td>
<td>6-9**(1)**</td>
<td>(+1)°</td>
<td>2</td>
<td>9-12**</td>
</tr>
<tr>
<td>IGHV5</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>IGHV6</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>IGHV7</td>
<td>0-1**(1)**</td>
<td>1</td>
<td>4</td>
<td>5-6**</td>
</tr>
<tr>
<td>IGHV(II)</td>
<td>-</td>
<td>-</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>IGHV(III)</td>
<td>-</td>
<td>-</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>IGHV(IV)</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>38-44**(2)**</td>
<td>4**(1)**</td>
<td>79**(2)**</td>
<td>123-129**</td>
</tr>
</tbody>
</table>

### IMGT Alignment of alleles

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Q</td>
<td>S</td>
<td>V</td>
<td>T</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>S</td>
<td>Q</td>
<td>V</td>
<td>P</td>
<td>V</td>
<td>F</td>
<td>E</td>
<td>A</td>
<td>P</td>
<td>V</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>GCC</td>
<td>CAG</td>
<td>CGG</td>
<td>TCT</td>
<td>GTG</td>
<td>ACC</td>
<td>CAG</td>
<td>CGG</td>
<td>TCT</td>
<td>ACC</td>
<td>CAG</td>
<td>GTG</td>
<td>CTT</td>
<td>GCC</td>
<td>CTT</td>
<td>ACC</td>
<td>CAG</td>
<td>GTG</td>
<td>CTT</td>
<td>ACC</td>
</tr>
</tbody>
</table>

**CDR1-IMGT**

<table>
<thead>
<tr>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>R</td>
<td>C</td>
<td>H</td>
<td>Y</td>
<td>S</td>
<td>S</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>Y</td>
<td>L</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTG</td>
<td>ACC</td>
<td>TGG</td>
<td>TAC</td>
<td>TAC</td>
<td>TCA</td>
<td>TGG</td>
<td>CTC</td>
<td>TTT</td>
<td>TCA</td>
<td>GTG</td>
<td>CTT</td>
<td>ACC</td>
<td>CAG</td>
<td>CTC</td>
<td>CTT</td>
<td>CAC</td>
<td>GTG</td>
<td>TAT</td>
<td>TAA</td>
</tr>
</tbody>
</table>

**CDR2-IMGT**

<table>
<thead>
<tr>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>46</th>
<th>47</th>
<th>48</th>
<th>49</th>
<th>50</th>
<th>51</th>
<th>52</th>
<th>53</th>
<th>54</th>
<th>55</th>
<th>56</th>
<th>57</th>
<th>58</th>
<th>59</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>Y</td>
<td>V</td>
<td>Q</td>
<td>Y</td>
<td>P</td>
<td>N</td>
<td>Q</td>
<td>G</td>
<td>L</td>
<td>Q</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>K</td>
<td>Y</td>
<td>L</td>
<td>S</td>
<td>O</td>
<td>S</td>
</tr>
<tr>
<td>GTG</td>
<td>TAT</td>
<td>GTG</td>
<td>CAA</td>
<td>TAC</td>
<td>CAC</td>
<td>TAC</td>
<td>TAC</td>
<td>CAC</td>
<td>TAC</td>
<td>CAG</td>
<td>CAG</td>
<td>CAC</td>
<td>TAC</td>
<td>CAC</td>
<td>TAC</td>
<td>CAC</td>
<td>TAC</td>
<td>TCA</td>
<td>GGA</td>
</tr>
</tbody>
</table>

**IMGT**

<table>
<thead>
<tr>
<th>61</th>
<th>62</th>
<th>63</th>
<th>64</th>
<th>65</th>
<th>66</th>
<th>67</th>
<th>68</th>
<th>69</th>
<th>70</th>
<th>71</th>
<th>72</th>
<th>73</th>
<th>74</th>
<th>75</th>
<th>76</th>
<th>77</th>
<th>78</th>
<th>79</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>L</td>
<td>L</td>
<td>E</td>
<td>S</td>
<td>I</td>
<td>N</td>
<td>V</td>
<td>E</td>
<td>S</td>
<td>I</td>
<td>N</td>
<td>V</td>
<td>E</td>
<td>S</td>
<td>I</td>
<td>N</td>
<td>V</td>
<td>E</td>
<td>S</td>
</tr>
<tr>
<td>ACC</td>
<td>CTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CDR3-IMGT**

<table>
<thead>
<tr>
<th>101</th>
<th>102</th>
<th>103</th>
<th>104</th>
<th>105</th>
<th>106</th>
<th>107</th>
<th>108</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>V</td>
<td>F</td>
<td>C</td>
<td>R</td>
<td>V</td>
<td>S</td>
<td>E</td>
</tr>
<tr>
<td>GAA</td>
<td>TAC</td>
<td>TTC</td>
<td>TGG</td>
<td>GCT</td>
<td>TGG</td>
<td>AGT</td>
<td>GA</td>
</tr>
</tbody>
</table>

**AE000653, TRAV0-6*02**

AE000653, TRAV0-6*02

<table>
<thead>
<tr>
<th>X02850</th>
<th>TRAV0-6*01</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE000653, TRAV0-6*02</td>
<td></td>
</tr>
<tr>
<td>M86361, TRAV0-6*02</td>
<td></td>
</tr>
</tbody>
</table>

## IMGT Protein display

### TRAV gene

<table>
<thead>
<tr>
<th>TRAV</th>
<th>FR1-IMGT (1-26)</th>
<th>CDR1-IMGT (27-38)</th>
<th>FR2-IMGT (39-55)</th>
<th>CDR2-IMGT (56-65)</th>
<th>FR3-IMGT (66-104)</th>
<th>CDR3-IMGT (105-115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE000658, TRAV1-1</td>
<td>QSLEQ PSEVTAVEQAIQIMCITYQ</td>
<td>TSGCYC</td>
<td>LSWYQQHDGAPFTLYS</td>
<td>KGDO</td>
<td>LEETG</td>
<td>RPSSLSRSGYVLLOELQKEDASAYFC AVR</td>
</tr>
<tr>
<td>AE000658, TRAV1-2</td>
<td>QNIDQ PTVTATEQSIIQIMCITYQ</td>
<td>TSGCYC</td>
<td>LPWVQHMAGEADPFTLYS</td>
<td>NVLDG</td>
<td>LEEMG</td>
<td>RPSSLSRSGYVLLKELQKEDASAYLC AVR</td>
</tr>
<tr>
<td>AE000658, TRAV2</td>
<td>KDQVQ PSTVAESEQAVFECQNSH</td>
<td>VSNAYN</td>
<td>FPWLYHPGCAAPRLLK</td>
<td>OQK</td>
<td>PSGQG</td>
<td>RNMTYER_PSSLLLQHEADAAVYTC AVL</td>
</tr>
<tr>
<td>AE000658, TRAV3</td>
<td>AGSYAQPEDQVQAEQCTCCTY</td>
<td>VSGHY</td>
<td>LPWVQYPQGRLQLQIL</td>
<td>YTGOML</td>
<td>YVGQY</td>
<td>GEAEFKQSGPLQPSLWSMDASTSYFC AVD</td>
</tr>
<tr>
<td>AE000658, TRAV4</td>
<td>LAKTQ PIKNDSYSEQEQVHITQSH</td>
<td>NLKNYD</td>
<td>ITWYQGPSQGRPPIQ</td>
<td>GYKT</td>
<td>KTVNE</td>
<td>WASLIPADREKLSVRPSLSDSAYFC AVD</td>
</tr>
<tr>
<td>AE000659, TRAV5</td>
<td>GEDVEQG LPSVEQADVKVIMCITYT</td>
<td>DSSSTY</td>
<td>LYTQKQEIPGQSLQLLTY</td>
<td>IFSGMQ</td>
<td>MKQQG</td>
<td>VVLYVLMKEDKSLRINADTSGAASAYFC AEL</td>
</tr>
<tr>
<td>AE000659, TRAV6</td>
<td>SQIEQNSALNQEGQTALGNTC</td>
<td>NYSAPY</td>
<td>LOWRQDPGRQGVPLL</td>
<td>IRENEK</td>
<td>EKREK</td>
<td>RKLVTFDTQKLGSHLHSGAPDASATLC AEL</td>
</tr>
<tr>
<td>AE000659, TRAV7</td>
<td>ENQVEHSPHLGQGODQAVSMCSOTY</td>
<td>VSNAYN</td>
<td>LQWYQNDOMQKHELLS</td>
<td>MYSAGY</td>
<td>EKQKC</td>
<td>RNWACKY_LQSSLYTQAVQPSDASTFAVYD</td>
</tr>
<tr>
<td>AE000659, TRAV8-1</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000659, TRAV8-2</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000659, TRAV8-3</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000659, TRAV8-4</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000659, TRAV8-5</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000659, TRAV8-6</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>Z02850, TRAV8-7</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
<tr>
<td>AE000660, TRAV8-7</td>
<td>AQVQOSQMVHULSEAELGQCNYS</td>
<td>YSGTVN</td>
<td>LPWVQYPQGQLQIL</td>
<td>YPGGDD</td>
<td>VGGIK</td>
<td>TPAEFITSKSFNLRSKQGQWSQDSTDAVFC AVN</td>
</tr>
</tbody>
</table>

IMGT Collier de Perles

IMGT databases and tools

The international ImMunoGeneTics information system®
Coordinator: M.-P. Lefranc, Montpellier, France http://imgt.cines.fr
Locus representation: Human IGL

Human IGL 22q11.21

WELCOME!
to IMGT/GENE-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

http://imgt.cines.fr
IMGT/LIGM-DB ON LINE, HERE YOU ARE!

Five types of search are available: select one by clicking on the button:

- accession number, mnemonic, definition, creation date, length, annotation level

**Catalogue**
cDNA (in databases: mRNA!)

...........gagga ttcaccatgg aactggggtct cggctgggttt ttccttgttg cttttttaga 120
aggtgtccag tgtgaggtgc aactggtgga gtcttggggga ggcctggctca acgccaagggg 180
gtccccggaga cttcctcttg cagcctcttg attaaagctc agtaacctag ctcatgagctg 240
ggtccggcagg gctccaggga agggagctga atgggtctca agtattagta gtagaaggtga 300
ttacatatac tatagagagt cagtgagggg ccagttcacc atctccagag acaacgccaa 360
gaattcacttg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgccgagagat tcttggtaatg tgctatatat gtatgtgatt cagctccctgg gccagggaac 480
cctggtcacc gttctctcag catccccgcag cagcccccaag gtttttccgc tgtgctctctg 540
cagcaccacc ccagacaggg aagttgagcc ctgctgccctg cctccctggg cctccctgag 600
ggagccacttc agtggtgacct ggagcgaaag cgggacccgc ccggtgcgtc ccggtgcgtcag 660
ccaggtgatgtg actgcttccc gggccagttc ccatcactcca cccaccccct cccaccccc 720
tccacctacc ccacctccct cctgcttgca cccggcagctg gccgagctcc gccgagctcc 780
cggagctgctg cggagctgctg cggagctgctg cggagctgctg cggagctgctg cggagctgctg 840
tccacctacc ccacccctcc ccagctccctg cgggagctcc gccgagctcc gccgagctcc 900
tgacctcctcag cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 960
tccacctacc ccacccctcc ccagctccctg cgggagctcc gccgagctcc gccgagctcc 1020
acccgtgggtc gctcccttcc gttgtgggca ccttttccgg gttgtgggca ccttttccgg 1080
gccatgagaac cccggcagg gctcccttcc gttgtgggca ccttttccgg gttgtgggca 1140
gcctccctcc cccggcagctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1200
gtccctccctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1260
gtccctccctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1320
gtccctccctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1380
gtccctccctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1440
gtccctccctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg cggctgtgctg 1500
taaaccacc catgtcataag tgtctggttgt catggcgggag gttggacgga cctgctctga 1560 ....
The IMGT-ONTOLOGY concepts

<table>
<thead>
<tr>
<th>ID</th>
<th>HSCOMPRAA IMGT/LIGM annotation : automatic;</th>
<th>RNA</th>
<th>HUM; 375 BP.</th>
</tr>
</thead>
<tbody>
<tr>
<td>XX</td>
<td>L39956;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DT</td>
<td>28-SEP-2001 (Rel. 200139-5, arrived in LIGM-DE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DT</td>
<td>13-NOV-2001 (Rel. 200146-2, Last updated, Version 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XX</td>
<td>Homo sapiens monospecific anti-ssDNA antibody heavy chain variable region mRNA, complementarity determining regions 1-3 and framework regions 1-4.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>RNA; rearranged configuration; Ig-Heavy; regular; functionality productive; group IGHV; subgroup HV3; specificity anti-DNA single-stranded (ss) [human].</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XX</td>
<td>antigen receptor; immunoglobulin superfamily; Ig; Ig-Heavy; variable; diversity; joining; hybridoma; immunoglobulin.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td>Homo sapiens [human]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>Cetartiodactyla; Ruminidae; Homo.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RN</td>
<td>[1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RP</td>
<td>1-375</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RX</td>
<td>MEDLINE: 961951558.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>&quot;V gene sequences of human anti-ssDNA antibodies secreted by lupus-derived CD5-negative B cell hybridomas&quot;;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XX</td>
<td>EMBL; L39956.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XX</td>
<td>Key: Location/Qualifiers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PH</td>
<td>L-Y-D-J-C-SEQUENCE &lt;1..375&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PH</td>
<td>partial</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The IMGT-ONTOLOGY concepts

The figure shows a table with several columns and rows, each containing information related to IMGT-ONTOLOGY concepts. The table is divided into sections labeled 'DESCRIPTION', 'OBTENTION', 'CLASSIFICATION', and 'NUMEROTATION'. Each section highlights different aspects of the IMGT-ONTOLOGY concepts, such as location/qualifiers, partial annotation, genomic context, and specific gene regions. The table contains details like taxonomic information, cell types, and sequence data, which are crucial for understanding the classification and obtention of these concepts.

For example, the 'DESCRIPTION' section includes information about the position of the sequence, with specific locations such as L-V-D-J-C-SEQUENCE and V-D-J-REGION. The 'OBTENTION' section might detail how the sequence was obtained, possibly through partial annotation or specific genomic context. The 'CLASSIFICATION' section could outline the taxonomic classification of the organism, such as Homo sapiens. The 'NUMEROTATION' section might focus on the specific numerical aspects of the sequence, such as the range of amino acids.
WELCOME!
to the IMGT/V-QUEST Search page

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®


Analyse your Immunoglobulin nucleotide sequences

- Human
- Mouse
- Chondrichthyes
- Teleostei (contains now 19 species)
  - Atlantic cod
  - Channel catfish
  - Rainbow trout
- Sheep

Analyse your T cell Receptor nucleotide sequences

- Human
- Mouse
- Non-human primates
Analyse your Immunoglobulin sequences

Your selection: Human

Your sequences are compared to the Human IG set from the IMGT/V-QUEST reference directory sets

- Include the analysis of the junction by IMGT/JunctionAnalysis
- Do not include the analysis of the junction by IMGT/JunctionAnalysis

Sequence

You may give your sequence a name (optional):

Type (or copy/paste) your sequence into the box below:

gaggtgcagccgtttgagctgggggggtctctgagaacc
tctctgtgacagccctgattaacctttacacagtgcctgtggcagctccgcaggt
ttgcctggagggcttggcaacttgaagatatagtggtgctgcctgttgataactaacaagc
gagctccctcgagggccgccccacatctgccacagacaacctcaagaatgacgtttt
tgcasaatgataacgccacagccagacacagccgtatattaacctttgtgcgagacaagttt

- Or give the path access to a local file containing your sequence:

Start  Clear the form
Sequence compared with the Human IG set from the IMGT reference directory

- Hyphens (-) show nucleotide identity, dots (.) represent gaps

Alignment for V-GENE

<table>
<thead>
<tr>
<th>input</th>
<th>IGHV3-23*01</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>M98560</td>
<td></td>
<td>125</td>
</tr>
<tr>
<td>J003206</td>
<td>IGHV3-23*02</td>
<td>1263</td>
</tr>
<tr>
<td>U20901</td>
<td>IGHV3-23*03</td>
<td>1209</td>
</tr>
<tr>
<td>212398</td>
<td>IGHV3-48*03</td>
<td>1128</td>
</tr>
<tr>
<td>M98575</td>
<td>IGHV3-48*01</td>
<td>1128</td>
</tr>
</tbody>
</table>

Alignment for D-GENE

<table>
<thead>
<tr>
<th>input</th>
<th>IGHD3-3*02</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z09618</td>
<td>IGHD3-3*02</td>
<td>61</td>
</tr>
<tr>
<td>X13972</td>
<td>IGHD3-3*03</td>
<td>61</td>
</tr>
<tr>
<td>X97051</td>
<td>IGHD6-19*01</td>
<td>34</td>
</tr>
<tr>
<td>X13972</td>
<td>IGHD3-9*01</td>
<td>34</td>
</tr>
<tr>
<td>J00254</td>
<td>IGHD2-15*01</td>
<td>34</td>
</tr>
</tbody>
</table>

Alignment for J-GENE

<table>
<thead>
<tr>
<th>input</th>
<th>IGHJ5*02</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z663355</td>
<td>IGHJ5*02</td>
<td>213</td>
</tr>
<tr>
<td>J00254</td>
<td>IGHJ5*01</td>
<td>201</td>
</tr>
<tr>
<td>Z663355</td>
<td>IGHJ6*02</td>
<td>177</td>
</tr>
</tbody>
</table>

Applet input [1]
Translation of the JUNCTION

```
input
TGT GCG AAC AGG TTT TGG AGT GCT TAT CTG TGG TTC GAC CCC GCC CAG GCA
```

```
104
CAKRFWYDGFPWKGQG
```

```
116
```
Alignment with **FR-IMGT** and **CDR-IMGT** delimitations

```
<------------------------ F R 1 - I R G T ------------------------

1
M99660 IGHV3-23*01
J00235 IGHV3-23*02
U29491 IGHV3-23*03
Z12358 IGHV3-48*03
M99675 IGHV3-48*01

5
GAG GTG CAG CTG TTG GAG TCG GGC GGA

10
GGC TTG GTA CAG CCT GGG GGG TCC

15

<------------------------> CDR1 - IMGT

20
input
M99660 IGHV3-23*01
J00235 IGHV3-23*02
U29491 IGHV3-23*03
Z12358 IGHV3-48*03
M99675 IGHV3-48*01

25
CTG AGA CTG TCC TGT GCA GCC TCT

30
GGA TTC ACC TTT AGC AAG TGT GGC

35

<------------------------ F R 2 - I M G T ------------------------

input
M99660 IGHV3-23*01
J00235 IGHV3-23*02
U29491 IGHV3-23*03
Z12358 IGHV3-48*03
M99675 IGHV3-48*01

40
ATG AGC TGG GTG GCC CAG GGT CTT GGG AAG GGC

45
GCT GCC AAG GTG GTC TCA

50

<------------------------>
```

[^Applet input [1]:](http://imgt.cines.fr)
Collier de Perles
Some improvements for the next release
Analysis of several sequences (in FASTA format) in a single run.
Case of complementary reverse sequences

Sequence number 1: test

Complementary reverse sequence compared with the human IG set from the IMGT reference directory

>test (complementary reverse)

```caggtgcagcttggtgagctgcttcggtgagccgcttcctccagcctgggaggttgcctttgagactc
tctctttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
caggcaagggccgttgaatggtgctccggttaagctggtttgatgtgacactcaaaatatctct
gcagagtcgccagaggccgattcaccctctccagacaaattccagaaacacatgtttttgtc
ttcgcaastgaaccgcgtcgagggcgcagcccgacgcgctgtgttcaggctctgctgcagagctc
gttggctattaagggggaacaggggtcctctgagcttggcggccagggaacaggtc
acgtctcttcagacagacggtcgtaattcctacgctctattctatgctactccggtccacggtc
```
Selection of output parameters
Alignment for V-GENE

Closest V-GENES (evaluated from the 1rst nt to the 2nd-CYS codon plus 15 nt of the CDR3-IMGT)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Score</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>X12686</td>
<td>1243</td>
<td>93.61%</td>
</tr>
<tr>
<td>X12687</td>
<td>1198</td>
<td>91.84%</td>
</tr>
<tr>
<td>L57729</td>
<td>1184</td>
<td>92.08%</td>
</tr>
<tr>
<td>X74460</td>
<td>1171</td>
<td>90.70%</td>
</tr>
<tr>
<td>X72820</td>
<td>1171</td>
<td>90.78%</td>
</tr>
</tbody>
</table>

Alignment

---

Score and nucleotide identity
Alignment for J-GENE

Closest J-GENEs

<table>
<thead>
<tr>
<th>J00242</th>
<th>IGKJ4*01</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score</td>
</tr>
<tr>
<td>152</td>
<td>94.11%</td>
</tr>
<tr>
<td>AF103571</td>
<td>IGKJ4*02</td>
</tr>
<tr>
<td>143</td>
<td>91.17%</td>
</tr>
<tr>
<td>J00242</td>
<td>IGKJ3*01</td>
</tr>
<tr>
<td>107</td>
<td>79.41%</td>
</tr>
<tr>
<td>270250</td>
<td>IGKJ2*02</td>
</tr>
<tr>
<td>107</td>
<td>79.41%</td>
</tr>
<tr>
<td>J00242</td>
<td>IGKJ2*01</td>
</tr>
</tbody>
</table>

Score and nucleotide identity
IMGT/JunctionAnalysis
V-DOMAINs: VH and V-KAPPA

VH

V-D-J junction

V-J junction

V-KAPPA

Side view of the V-DOMAINs

View from above the CDRs

Mouse (Mus musculus) E5.2Fv

CDR3-IMGT = Complementarity determining region (105-117)
V-J junction (104-118)
V-D-J junction (104-118)
Immunoglobulin V-D-J generation of sequence diversity

3′V-REGION  N-REGION  D-REGION  N-REGION  5′J-REGION

tgtgcgaaa ga tac agcatatttg tgtgtgactgctat tcc gat gca acaacctggttcg actcctgg

JUNCTION

CAPYRGDTYDSW
tgt gcc cca tac cgg ggt gac act tat gat tac tcc tgg
Analysis of the JUNCTIONs

<table>
<thead>
<tr>
<th>Input</th>
<th>V name</th>
<th>V-REGION</th>
<th>N1</th>
<th>D-REGION</th>
<th>N2</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 M62724</td>
<td>IGHV7-4-1*02</td>
<td>ttagcagaggaagaagaagaagata ..</td>
<td></td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>#2 247269</td>
<td>IGHV1-69*06</td>
<td>ttagcagaggaagggggttaaggg ..</td>
<td>..</td>
<td>..</td>
<td>tcaggtgggtt</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input</th>
<th>J-REGION</th>
<th>J name</th>
<th>D name</th>
<th>Vmut</th>
<th>Dmut</th>
<th>Jmut</th>
<th>Nc</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 M62724</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>#2 247269</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
</tbody>
</table>

Translation of the JUNCTIONs

```
<table>
<thead>
<tr>
<th>105</th>
<th>106</th>
<th>107</th>
<th>108</th>
<th>110</th>
<th>111</th>
<th>112.4</th>
<th>112.2</th>
<th>112.1</th>
<th>113</th>
<th>114</th>
<th>116</th>
<th>118</th>
<th>CDR3-IMGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>A</td>
<td>R</td>
<td>E</td>
<td>D</td>
<td>S</td>
<td>N</td>
<td>G</td>
<td>Y</td>
<td>K</td>
<td>I</td>
<td>P</td>
<td>D</td>
<td>Y</td>
</tr>
<tr>
<td>#1 M62724 tgg ggc aga gaa gat agc aat ggc</td>
<td>tac aaatataaattgtacacatgcttggg + 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#2 247269 tgg ggc aga gaa ggg ggt gct aag gtc gaa ttt ttg gag tgg ttt gat ggg</td>
<td>tac tgg ttc gac ccc tgg + 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

--> IMGT/JunctionAnalysis Search page
--> IMGT/JunctionAnalysis Documentation
THANK YOU for using **IMGТ/JunctionAnalysis**

THE INTERNATIONAL IMMUNOGENETICS INFORMATION SYSTEM®

---

**Analysis of the JUNCTIONs**

<table>
<thead>
<tr>
<th>Input</th>
<th>V name</th>
<th>V-REGION</th>
<th>N1</th>
<th>D-REGION</th>
<th>N1</th>
<th>J-REGION</th>
<th>J name</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 962724</td>
<td>IGKV7-4-1*02</td>
<td>tgcggagaga</td>
<td>aga</td>
<td>tagcaatggctacaa</td>
<td>...</td>
<td>mcca</td>
<td>......</td>
</tr>
<tr>
<td>#2 247269</td>
<td>IGKV1-69*06</td>
<td>ttcggaga</td>
<td>dggccagctacaa</td>
<td>...</td>
<td>tcgcattttggctactggt</td>
<td>.....</td>
<td>tcaagttag</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Input</th>
<th>D name</th>
<th>Vsub</th>
<th>Dsub</th>
<th>Jmut</th>
<th>Nuc</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 962724</td>
<td>IGHDS-24*01</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1/7</td>
</tr>
<tr>
<td>#5 247269</td>
<td>IGHD3-3*02</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>14/21</td>
</tr>
</tbody>
</table>

---

**Translation of the JUNCTIONs**

<table>
<thead>
<tr>
<th>105</th>
<th>107</th>
<th>109</th>
<th>111</th>
<th>111.2</th>
<th>112.4</th>
<th>112.2</th>
<th>112</th>
<th>114</th>
<th>116</th>
<th>118</th>
<th>CDR3-IMGТ</th>
</tr>
</thead>
<tbody>
<tr>
<td>104</td>
<td>106</td>
<td>108</td>
<td>110</td>
<td>111</td>
<td>111.1</td>
<td>111.2</td>
<td>112.4</td>
<td>112.2</td>
<td>112</td>
<td>114</td>
<td>116</td>
</tr>
</tbody>
</table>

CAR EDS N G

| 1 | 962724 | tgc gga gaa gaa gat arc sat gge | tac aaa ata ttt gac tac tgg | + | 13 |
| 2 | 247269 | tgc gga gaa ggg ggg gtt cag ggg ttt tgg tgg ttc cat ggg tac tgg tgg | + | 20 |

Terminé
### IMGT/JunctionAnalysis Results

**Locus**: IGH  
**Species**: Homo sapiens  
**IMGTreertoire link**: Locus representation  

**Description of the JUNCTIONs**

Click on mutated (underscored) nucleotid to see the original one.

<table>
<thead>
<tr>
<th>Input</th>
<th>V name</th>
<th>V-REGION</th>
<th>P</th>
<th>N1</th>
<th>D-REGION</th>
<th>N2</th>
<th>P</th>
<th>J-REGION</th>
<th>J name</th>
<th>D name</th>
<th>Vant</th>
<th>Dmut</th>
<th>Jmut</th>
<th>Npc</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Z70256</td>
<td>IGTV2-2*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>Z70257</td>
<td>IGTV2-7*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td>Z70606</td>
<td>IGTV4-31*03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>Z70608</td>
<td>IGTV4-59*05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td>Z70610</td>
<td>IGTV4-34*08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>Z70611</td>
<td>IGTV4-59*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td>Z70613</td>
<td>IGTV4-59*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>Z70614</td>
<td>IGTV4-59*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>Z70616</td>
<td>IGTV4-59*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#10</td>
<td>Z70616</td>
<td>IGTV4-34*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#11</td>
<td>Z70620</td>
<td>IGTV4-30*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#12</td>
<td>Z70621</td>
<td>IGTV4-30*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#13</td>
<td>Z70622</td>
<td>IGTV4-39*06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The eleven IMGT amino acid classes according to the physico-chemical properties

JUNCTION alignments

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

<table>
<thead>
<tr>
<th>Frame</th>
<th>CDR1-MHG1 length</th>
<th>Molecular mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 Z70256</td>
<td>15</td>
<td>2,456.76</td>
</tr>
<tr>
<td>#2 Z70257</td>
<td>16</td>
<td>2,056.49</td>
</tr>
<tr>
<td>#3 Z70808</td>
<td>11</td>
<td>1,624.77</td>
</tr>
<tr>
<td>#4 Z70809</td>
<td>18</td>
<td>2,426.98</td>
</tr>
<tr>
<td>#5 Z70810</td>
<td>19</td>
<td>2,539.73</td>
</tr>
<tr>
<td>#6 Z70811</td>
<td>15</td>
<td>2,072.21</td>
</tr>
<tr>
<td>#7 Z70813</td>
<td>12</td>
<td>1,551.71</td>
</tr>
<tr>
<td>#8 Z70814</td>
<td>13</td>
<td>1,846.01</td>
</tr>
<tr>
<td>#9 Z70815</td>
<td>14</td>
<td>2,419.75</td>
</tr>
<tr>
<td>#10 Z70816</td>
<td>12</td>
<td>1,602.0</td>
</tr>
<tr>
<td>#11 Z70820</td>
<td>14</td>
<td>1,932.17</td>
</tr>
<tr>
<td>#12 Z70821</td>
<td>16</td>
<td>2,207.69</td>
</tr>
</tbody>
</table>
The international ImMunoGeneTics information system®
Coordinator: M.-P. Lefranc, Montpellier, France   http://imgt.cines.fr
WELCOME!

to IMGT/3Dstructure-DB

THE INTERNATIONAL IMMUNOGENETICS INFORMATION SYSTEM®

TRB V-D-J REGION

TRA V-J REGION

http://imgt.cines.fr
### IMGT/3Dstructure-DB card for: 1u8k

<table>
<thead>
<tr>
<th>Chain ID</th>
<th>C-REGION</th>
<th>V-DOMAIN</th>
<th>D-REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1u8k_A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1u8k_B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1u8k_C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Chain ID: 1u8k_A
- **IMGT chain description:** L-KAPPA
- **IMGT domain description:**
  - IGKV*1-1 (99.98%)
  - IGKV*1-2 (99.98%)
  - IGKV*1-4 (99.98%)
  - IGKV*3-1 (99.98%)
  - IGKV*4-1 (99.98%)
  - IGKV*4-2 (99.98%)
- **2D representation:** Collier de Perles or Collier de Perles on 2 layers
- **Sheet composition:** [A B D E] [C C' C' F G]

#### Chain ID: 1u8k_B
- **IMGT chain description:** VH-CH1
- **IMGT domain description:**
  - IGHV2-5*25 (55.00%)
  - IGHV2-5*6 (55.00%)
- **2D representation:** Collier de Perles or Collier de Perles on 2 layers
- **Sheet composition:** [A B D E] [C C' D' E']

#### Chain ID: 1u8k_C
- **IMGT chain description:** CH1
- **IMGT domain description:**
  - IGHG*1-2 (97.94%)
  - IGHG*1-3 (97.94%)
  - IGHG*1-4 (97.94%)
  - IGHG*5-1 (97.94%)
  - IGHG*5-2 (97.94%)
- **2D representation:** Collier de Perles or Collier de Perles on 2 layers
- **Sheet composition:** [A B D E] [C F G]

#### Experiment details
- **Protein name:** 2FS
- **Receptor description:** FAB-GAMMA-3_KAPPA
- **Type:** IG
- **Ligand(s):** GP41
- **Species:** Homo sapiens (Human)
- **CC:** 1
- **Resolution:** 2.24
- **PDB release date:** 05-OCT-04
- **Experimental technique:** X-ray diffraction

---

### References
- [IMGT/3Dstructure-DB](http://imgt.cines.fr)
<table>
<thead>
<tr>
<th><strong>Chain ID</strong></th>
<th><strong>1u8k_B</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IMGT chain description</strong></td>
<td><strong>VH-CH1</strong></td>
</tr>
<tr>
<td><strong>Region identification and delimitation</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Chain amino acid sequence</strong></th>
</tr>
</thead>
</table>
| [IG

| **IMGT domain description** | **VH** |
| **IGHV gene and allele name** | IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%), [Alignment details](#) |
| **IGHJ gene and allele name** | IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), [Alignment details](#) |
| **2D representation** | Collier de Perles or Collier de Perles on 2 layers |
| **Sheet composition** | [A B D E] [C C" C' F G] |
| | [CDR1] [CDR2] |
| | RITLKEGSLV/KPTQLTLTCSFSFGSLSIDFGV/GVWIRQP/PGKALEWIAIYSDDK...RYSPSIN.TRLTIKDTSINKVVLV |
| | MTRVSPVD/DTATYFC/AR/HRGPPTT/LFGV/P1ARGP/ANMD/WQGI/TVI |

| **IMGT domain description** | **CH1** |
| **IGHC gene and allele name** | IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%), IGHG3*14 (97.94%), IGHG3*15 (97.94%), IGHG3*16 (97.94%), IGHG3*17 (97.94%), IGHG3*20 (97.94%), [Alignment details](#) |
| **2D representation** | Collier de Perles or Collier de Perles on 2 layers |
| **Sheet composition** | [A B D E] [C F G] |
| | .....STKGPSV/FP/A/PCS/KSTA...G/G/TAALGCLV/KDYFP..EP/TVSW/NSGALTS...GV/HTFP/AVLQS/SS...GLYSLSVVTV/PSSSL...GTQ/TQVN/NHKPSNTK/V/DRKV |
Identification of the IMGT gene and the IMGT closest allele(s)

<table>
<thead>
<tr>
<th>Chain ID</th>
<th>1u8k_B</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMGT chain description</td>
<td>VH-CH1</td>
</tr>
<tr>
<td>Chain amino acid sequence</td>
<td>[IGHV1 RITLKE SGPLV KPTQ TLTLTCSFSGFSL DFGV GWRIRQPPGALE WIAIYSDDKRYS PSINTRLTI TKDS NKQVVLVMTRVSPV ] [ D ] [ IGHJ1 ] [ IGDTATYFC AHRGPTTLFGV PIARGPNAMD WQQGITVTI ]</td>
</tr>
</tbody>
</table>

**V-DOMAIN**

<table>
<thead>
<tr>
<th>IMGT domain description</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGHV gene and allele name</td>
<td>IGHV2-5<em>05 (85.00%), IGHV2-5</em>06 (85.00%), [Alignment details]</td>
</tr>
<tr>
<td>IGHJ gene and allele name</td>
<td>IGHJ6<em>01 (83.33%), IGHJ6</em>02 (83.33%), [Alignment details]</td>
</tr>
</tbody>
</table>

**2D representation**
- Collier de Perles or Collier de Perles on 2 layers

**CDR-IMGT lengths**
- [10.7.24]

**Sheet composition**
- [A B D E] [C C' F G]

**IMGT domain description**
- CH1

**IGHC gene and allele name**
- IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%), IGHG3*14 (97.94%), IGHG3*15 (97.94%), IGHG3*16 (97.94%), IGHG3*17 (97.94%), IGHG3*20 (97.94%), [Alignment details]

**2D representation**
- Collier de Perles or Collier de Perles on 2 layers

**Sheet composition**
- [A B D E] [C F G]

| [STKGPSVFPLAPCSKSTA...GCTAALGCLVKDYFP...EP/TVSWSNAGLTS...GVHTFPAVLQSS....GLYSLSSVVTVPSSSL...GTYTQNVNHKP...SNTKVDKRV] |
### Description of domains

**Chain ID**: VH-CH1

**IMGT domain description**: VH

**IGHV gene and allele name**: IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%), [Alignment details](#)

**IGHJ gene and allele name**:IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), [Alignment details](#)

**2D representation**: Collier de Perles or [Collier de Perles on 2 layers](#)

**CDR-IMGT lengths**: [10.724]

**Sheet composition**: [A B D E] [C C' F G]

```python
[        CDR1          ] [        CDR2          ]
RlTlKeSgp.PLvKPtqTlTlCfpsGlSlDfGv...Vg#IrqPpgkAlewLaiYsDdkd...Rysps1n.TrLtiTkDtsKnQvVlv
Cdr3]
MTrv sv/PdtatYFcAhRgPtTlFgv/Piargp/Namd/WQqGItVti
```

### CDR-IMGT lengths

**Gaps according to IMGT numbering**

**IMGT domain description**: CH1

**IGHC gene and allele name**: IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%), IGHG3*14 (97.94%), IGHG3*15 (97.94%), IGHG3*16 (97.94%), IGHG3*17 (97.94%), IGHG3*20 (97.94%), [Alignment details](#)

**2D representation**: Collier de Perles or [Collier de Perles on 2 layers](#)

**Sheet composition**: [A B D E] [C F G]

```python
.....StkGpsVfPlApCsKsta...GttaAlGClVkdYfp...Ep/TVswNsgalt...Gv/HtFpAvLqS...GlySlsSvVtv
Pssl...GtQtYtQnVnhkp...Sntk/DkrV
```
### Contact analysis

**41V - TRP (W)**

chain : 1u8k_B

<table>
<thead>
<tr>
<th>IMGT Num</th>
<th>Code 1C</th>
<th>Code 3C</th>
<th>Chain</th>
<th>Domain</th>
<th>Tot</th>
<th>NCo</th>
<th>Pol</th>
<th>HB</th>
<th>NPol</th>
<th>Cov</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>6V1</td>
<td>E</td>
<td>GLU</td>
<td>1u8k_B</td>
<td>VH</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>21V1</td>
<td>L</td>
<td>LEU</td>
<td>1u8k_B</td>
<td>VH</td>
<td>17</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>22V1</td>
<td>T</td>
<td>THR</td>
<td>1u8k_B</td>
<td>VH</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>23V1</td>
<td>C</td>
<td>CYS</td>
<td>1u8k_B</td>
<td>VH</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>39V1</td>
<td>V</td>
<td>VAL</td>
<td>1u8k_B</td>
<td>VH</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>43V1</td>
<td>R</td>
<td>ARG</td>
<td>1u8k_B</td>
<td>VH</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>51V1</td>
<td>E</td>
<td>GLU</td>
<td>1u8k_B</td>
<td>VH</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>52V1</td>
<td>W</td>
<td>TRP</td>
<td>1u8k_B</td>
<td>VH</td>
<td>11</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>53V1</td>
<td>L</td>
<td>LEU</td>
<td>1u8k_B</td>
<td>VH</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>54V1</td>
<td>A</td>
<td>ALA</td>
<td>1u8k_B</td>
<td>VH</td>
<td>18</td>
<td>18</td>
<td>4</td>
<td>2</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>55V1</td>
<td>I</td>
<td>ILE</td>
<td>1u8k_B</td>
<td>VH</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>78V1</td>
<td>I</td>
<td>ILE</td>
<td>1u8k_B</td>
<td>VH</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>87V1</td>
<td>V</td>
<td>VAL</td>
<td>1u8k_B</td>
<td>VH</td>
<td>11</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>88V1</td>
<td>V</td>
<td>VAL</td>
<td>1u8k_B</td>
<td>VH</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>89V1</td>
<td>L</td>
<td>LEU</td>
<td>1u8k_B</td>
<td>VH</td>
<td>33</td>
<td>33</td>
<td>1</td>
<td>0</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>102V1</td>
<td>Y</td>
<td>TYR</td>
<td>1u8k_B</td>
<td>VH</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>103V1</td>
<td>F</td>
<td>PHE</td>
<td>1u8k_B</td>
<td>VH</td>
<td>18</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>104V1</td>
<td>C</td>
<td>CYS</td>
<td>1u8k_B</td>
<td>VH</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Info</td>
<td>105V1</td>
<td>A</td>
<td>ALA</td>
<td>1u8k_B</td>
<td>VH</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Tot**  Total number of atomic pair contacts

**NCo**  Number of non covalent atomic contacts

**Pol**  Number of polar atomic pair contacts

**HB**   Number of hydrogen bonds

**NPol** Number of non polar atomic pair contacts

**Cov**  Number of covalent links (other than chain covalent links)

**SS**   Number of disulfide bridges
Collier de Perles: HUMAN IGHV V-DOMAIN from B12 (PDB: 1hzh_H)

[8.8.20]
IMGT, the international ImMunoGeneTics information system®

Created in 1989 (Université Montpellier II and CNRS)

1) IMGT domains of research:
   Immunogenetics - Immunoinformatics
   * Immunoglobulins (IG), T cell receptors (TR), major histocompatibility complex (MHC) of human and other vertebrates,
   * Immunoglobulin superfamily IgSF and MhcSF,
   * Related proteins of the immune system (RPI)

2) IMGT-ONTOLOGY concepts

3) IMGT-Choreography: Dynamic knowledge management
Structural domains

IG and TR

V-DOMAIN

C-DOMAIN

MHC

G-DOMAINs

http://imgt.cines.fr
Interactions between domains

Ch 2 (C-like domain) and [D2] C-like-domain interact

IGHG1 (FC-GAMMA1)

FCGR3B

http://imgt.cines.fr
MHC class I IMGT contact sites

H2-K1*01
(code 1jtr)
8 residue peptide

(Kaas et al. In press)
**MHC class I IMGT contact sites**

H2-K1*01 8 residue peptide (code 1jtr)
IMGT, the international ImMunoGeneTics information system®

http://imgt.cines.fr

The IMGT team at Montpellier, France