LaTeX 项目模板

Roger Young

2017 年 9 月 19 日

目录

第一章　General

Since references to web sites are not yet acknowledged as citations, please mention Den Dunnen et al. (2016) HGVS recommendations for the description of sequence variants: 2016 update. Hum.Mutat. 25: 37: 564-569 when referring to these pages. Note that although the examples on these pages mainly give examples for human (Homo sapiens), the recommendations can be applied to all species.

1.1 General recommendations

1. all variants should be described at the most basic level, **the DNA level**. Descriptions at the RNA and/or protein level may be given in addition.

   - descriptions should make clear whether the change was **experimentally determined or theoretically deduced** by giving predicted consequences in parentheses
   - descriptions at RNA/protein level should describe the changes observed on that level (RNA/protein) and not try to incorporate any knowledge regarding the change at DNA-level (see Questions below)

2. all variants should be described in relation to an **accepted reference sequence** (see Reference Sequences).

   - the reference sequence file used should be **public** and **clearly described**, e.g. `NC_000023.10`, `LRG_199`, `NG_012232.1`, ENST00000357033, `NM_004006.2`, `NR_002196.1`, `NP_003997.1`, etc.
     - when variants are not reported in relation to a genomic reference sequence from a recent genome build, the preferred reference sequence is a **Locus Reference Genomic sequence (LRG)**
     - when no LRG is available, one should be requested (see Reference Sequences).
     - the reference sequence used must contain the residue(s) described to be changed.
   - a letter prefix should be used to indicate the type of reference sequence used. Accepted prefixes are;
     - "g." for a **genomic reference sequence**
     - "c." for a **coding DNA reference sequence**
     - "n." for a **non-coding DNA reference sequence**

  - "r." for an **RNA reference sequence (transcript)**
  - "p." for a **protein reference sequence**
- numbering of the residues (nucleotide or amino acid) in relation to the reference sequence used should follow the approved scheme (see Numbering)

3. 3' rule: for all descriptions the most 3' position possible of the reference sequence is arbitrarily assigned to have been changed
   - the 3' rule also applies for changes in single residue stretches and tandem repeats (nucleotide or amino acid)
   - the 3' rule applies to ALL descriptions (genome, gene, transcript and protein) of a given variant
4. descriptions at DNA, RNA and protein level are clearly different:
   - **DNA-level** 123456A>T (see Details): number(s) referring to the nucleotide(s) affected, nucleotides in **CAPITALS** using **IUPAC-IUBMB assigned nucleotide symbols**
   - **RNA-level** 76a>u (see Details): number(s) referring to the nucleotide(s) affected, nucleotides in **lower case** using **IUPAC-IUBMB assigned nucleotide symbols**
   - **protein level** Lys76Asn (see Details): the amino acid(s) affected in **3- or 1-letter** followed by a number IUPAC-IUBMB assigned amino acid symbols * **three-letter amino acid code is preferred** (see Standards)
5. prioritisation: when a description is possible according to several types, the preferred description is: (1) deletion, (2) inversion, (3) duplication, (4) conversion, (5) insertion
   - when a variant can be described as a duplication or an insertion, prioritisation determines it should be described as a duplication
6. only approved **HGNC gene symbols** should be used to describe genes or proteins

1.2 Characters Used

In HGVS nomenclature some characters have a specific meaning

- "+" (plus) is used in nucleotide numbering; c.123+45A>G
- "-" (minus) is used in nucleotide numbering; c.124-56C>T
- "*" (asterisk) is used in nucleotide numbering and to **indicate a translation termination (stop) codon** (see Standards); c.*32G>A and P.Trp41*
- "_" (underscore) is used to **indicate a range**; `g.12345_12678del`
- "[ ]" (angled brackets) are used for **alleles** (see DNA, RNA, protein)

- ";" (semi colon) is used to **separate variants and alleles**; g.[123456A>G;345678G>C] or g.[123456A>G];[345678G>C]
- "," (comma) is used to separate different transcripts/proteins derived from one allele; `r.[123a>t, 122_154del]`
- ":" (colon) is used to separate the reference sequence file identifier (accession.version_number) from the actual description of a variant; `NC_000011.9:g.12345611G>A`
- "( )" (parentheses) are used to indicate uncertainties and predicted consequences; `NC_000023.9:g.(123456_23456` `p.(Ser123Arg)`

  NOTE: the range of the uncertainty should be described as precisely as possible (see below)
- "?" (question mark) is used to indicate unknown positions (nucleotide or amino acid); `g.(?_234567)_(345678_?)`
- "^" (caret) is used as "or" ; `c.(370A>C^372C>R)` as back translation of p.Ser124Arg
- ">" (greater than) is used to describe substitution variants (DNA and RNA level); `g.12345A>T`, `r.123a>u` (see DNA, RNA)
- "{ }" (curly braces) suggested for the description of variants in otherwise perfect copy sequences (see Open Issues); `g.24_65dup{46G>T}`
- "=" (equals) is used to indicate a sequence was tested but found unchanged; `p.(Arg234=)`
- "/" (forward slash) is used to indicate mosaicism (see Complex (HGVS/ISCN))
- "//" (double forward slash) is used to indicate chimerism (see Complex (HGVS/ISCN))

第一部分

附录

1.3 元素周期表

元 素 周 期 表

1 2.20 1s																	2 1s
H 氢 Hydrogen 1.00784–1.00811																	**He** 氖 Helium 4.002602(2)
3 0.98 2s	4 1.57 2s											5 2.04 2p	6 2.55 2p	7 3.04 2p	8 3.44 2p	9 3.98 2p	10 2p
Li 锂 Lithium 6.938–6.997	**Be** 铍 Beryllium 9.012183 1(5)											**B** 硼 Boron 10.806–10.821	**C** 碳 Carbon 12.0096–12.0116	**N** 氮 Nitrogen 14.00643–14.00728	**O** 氧 Oxygen 15.99903–15.99977	**F** 氟 Fluorine 18.998403163(6)	**Ne** 氦 Neon 20.1797(6)
11 0.93 3s	12 1.31 3s											13 1.61 3p	14 1.90 3p	15 2.19 3p	16 2.58 3p	17 3.16 3p	18 3p
Na 钠 Sodium 22.98976928(2)	**Mg** 镁 Magnesium 24.304–24.307											**Al** 铝 Aluminium 26.9815385(7)	**Si** 硅 Silicon 28.084–28.086	**P** 磷 Phosphorus 30.973761998(5)	**S** 硫 Sulphur 32.059–32.076	**Cl** 氯 Chlorine 35.446–35.457	**Ar** 氩 Argon 39.948(1)
19 0.82 4s	20 1.00 4s	21 1.36 3d	22 1.54 3d	23 1.63 3d	24 1.66 3d*	25 1.55 3d	26 1.83 3d	27 1.88 3d	28 1.91 3d	29 1.90 3d*	30 1.65 3d	31 1.81 4p	32 2.01 4p	33 2.18 4p	34 2.55 4p	35 2.96 4p	36 3.00 4p
K 钾 Potassium 39.0983(1)	**Ca** 钙 Calcium 40.078(4)	**Sc** 钪 Scandium 44.955908(5)	**Ti** 钛 Titanium 47.867(1)	**V** 钒 Vanadium 50.9415(1)	**Cr** 铬 Chromium 51.9961(6)	**Mn** 锰 Manganese 54.938044(3)	**Fe** 铁 Iron 55.845(2)	**Co** 钴 Cobalt 58.933194(4)	**Ni** 镍 Nickel 58.6934(4)	**Cu** 铜 Copper 63.546(3)	**Zn** 锌 Zinc 65.38(2)	**Ga** 镓 Gallium 69.723(1)	**Ge** 锗 Germanium 72.630(8)	**As** 砷 Arsenic 74.921595(6)	**Se** 硒 Selenium 78.971(8)	**Br** 溴 Bromine 79.901–79.907	**Kr** 氪 Krypton 83.798(2)
37 0.82 5s	38 0.95 5s	39 1.22 4d	40 1.33 4d	41 1.6 4d*	42 2.16 4d*	43 1.9 4d	44 2.2 4d*	45 2.28 4d*	46 2.20 4d*	47 1.93 4d*	48 1.69 4d	49 1.78 5p	50 1.96 5p	51 2.05 5p	52 2.1 5p	53 2.66 5p	54 2.60 5p
Rb 铷 Rubidium 85.4678(3)	**Sr** 锶 Strontium 87.62(1)	**Y** 钇 Yttrium 88.90584(2)	**Zr** 锆 Zirconium 91.224(2)	**Nb** 铌 Niobium 92.90637(2)	**Mo** 钼 Molybdenum 95.95(1)	**Tc** 锝 Technetium (98)	**Ru** 钌 Ruthenium 101.07(2)	**Rh** 铑 Rhodium 102.90550(2)	**Pd** 钯 Palladium 106.42(1)	**Ag** 银 Silver 107.8682(2)	**Cd** 镉 Cadmium 112.414(4)	**In** 铟 Indium 114.818(1)	**Sn** 锡 Tin 118.710(7)	**Sb** 锑 Antimony 121.760(1)	**Te** 碲 Tellurium 127.60(3)	**I** 碘 Iodine 126.90447(3)	**Xe** 氙 Xenon 131.293(6)
55 0.79 6s	56 0.89 6s	57-71	72 1.3 5d	73 1.5 5d	74 2.36 5d	75 1.9 5d	76 2.2 5d	77 2.20 5d	78 2.28 5d*	79 2.54 5d*	80 2.00 5d	81 1.62 6p	82 1.87 6p	83 2.02 6p	84 2.0 6p	85 2.2 6p	86 2.2 6p
Cs 铯 Cesium 132.9054 5196(6)	**Ba** 钡 Barium 137.327(7)	镧系 Lanthanides	**Hf** 铪 Halfnium 178.49(2)	**Ta** 钽 Tantalum 180.94788(2)	**W** 钨 Tungsten 183.84(1)	**Re** 铼 Rhenium 186.207(1)	**Os** 锇 Osmium 190.23(3)	**Ir** 铱 Iridium 192.217(3)	**Pt** 铂 Platinum 195.084(9)	**Au** 金 Gold 196.966569(5)	**Hg** 汞 Mercury 200.592(3)	**Tl** 铊 Thallium 204.382–204.385	**Pb** 铅 Lead 207.2(1)	**Bi** 铋 Bismuth 208.98040(1)	**Po** 钋 Polonium (209)	**At** 砹 Astatine (210)	**Rn** 氡 Radon (222)
87 0.7 7s	88 0.9 7s	89-103	104 6d	105 6d	106 6d	107 6d	108 6d	109 6d	110 6d	111 6d	112 6d	113 7p	114 7p	115 7p	116 7p	117 7p	118 7p
Fr 钫 Francium (223)	**Ra** 镭 Radium (226)	锕系 Actinides	**Rf** 𨧀 Rutherfordium (261)	**Db** □ Dubnium (268)	**Sg** □ Seaborgium (269)	**Bh** □ Bohrium (270)	**Hs** □ Hassium (269)	**Mt** 鿏 Meitnerium (278)	**Ds** 鐽 Darmstadtium (281)	**Rg** 錀 Roentgenium (282)	**Cn** 鿔 Copernicium (285)	**Nh** 鉨 Nihonium (286)	**Fl** 鈇 Flerovium (289)	**Mc** 镆 Moscovium (289)	**Lv** 鉝 Livermorium (293)	**Ts** Tennessine (294)	**Og** Oganesson (294)

镧系	57 1.1 5d*	58 1.12 4f*	59 1.13 4f	60 1.14 4f	61 1.13 4f	62 1.17 4f	63 1.2 4f	64 1.2 4f*	65 1.1 4f	66 1.22 4f	67 1.23 4f	68 1.24 4f	69 1.25 4f	70 1.1 4f	71 1.27 4f
	La 镧 Lanthanum 138.90547(7)	**Ce** 铈 Cerium 140.116(1)	**Pr** 镨 Praseodymium 140.90766(2)	**Nd** 钕 Neodymium 144.242(3)	**Pm** 钷 Promethium (145)	**Sm** 钐 Samarium 150.36(2)	**Eu** 铕 Europium 151.964(1)	**Gd** 钆 Gadolinium 157.25(3)	**Tb** 铽 Terbium 158.92535(2)	**Dy** 镝 Dysprosium 162.500(1)	**Ho** 钬 Holmium 164.93033(2)	**Er** 铒 Erbium 167.259(3)	**Tm** 铥 Thulium 168.93422(2)	**Yb** 镱 Ytterbium 173.045(10)	**Lu** 镥 Lutetium 174.9668(1)
锕系	89 1.1 6d*	90 1.3 5f*	91 1.5 5f*	92 1.38 5f*	93 1.36 5f*	94 1.28 5f	95 1.13 5f	96 1.28 5f*	97 1.3 5f	98 1.3 5f	99 1.3 5f	100 1.3 5f	101 1.3 5f	102 1.3 5f	103 1.3 5f
	Ac 锕 Actinium (227)	**Th** 钍 Thorium 232.0377(4)	**Pa** 镤 Protactinium 231.03588(2)	**U** 铀 Uranium 238.02891(3)	**Np** 镎 Neptunium (237)	**Pu** 钚 Plutonium (244)	**Am** 镅 Americium (243)	**Cm** 锔 Curium (247)	**Bk** 锫 Berkelium (247)	**Cf** 锎 Californium (251)	**Es** 锿 Einsteinium (252)	**Fm** 镄 Fermium (257)	**Md** 钔 Mendelevium (258)	**No** 锘 Nobelium (259)	**Lr** 铹 Lawrencium (266)

Z = atomic number; eneg = electronegativity; ss = subshell; Sy = Symbol, Name = element name, saw = standard atomic weight

Z eneg ss
Sy
Name
saw

相对原子质量来源：(http://ciaaw.org/atomic-weights.htm). . © 2017 张洋

An asterisk (*) next to a subshell indicates an anomalous (Aufbau rule-breaking) ground state electron configuration.