# Sequence Notation for SNPs Housed in dbSNP

Created: July 23, 2005; Updated: February 18, 2014.

**Define the term "orientation" as used in dbSNP.**

Submissions to our database have arbitrary orientation relative to each other. If multiple submissions refer to the same SNP, they may cluster together in reverse orientation, so we also track the orientation of each submission relative to the exemplar ss. Please bear in mind that submitters to dbSNP are only required to provide some flanking sequence around the SNP for context. The SNPdev team does the positioning using BLAST and the resulting alignments. **(3/13/05)**

**What does N/N represent in refSNP sequence?**

N/N is the IUPAC code used to indicate that the actual base can't be determined by a genotyping assay. **(11/17/05)**

**We have found lowercase (small) letters in SNP sequence reports downloaded from your FTP site. What do these lowercase letters mean?**

Lowercase (small) lettering is used for sequences identified by RepeatMasker as repetitive elements or as low complexity. You can find this information by accessing the refSNP page FASTA section and clicking on Legend.

**In SNP flank sequence, I find that some bases are capital letters, while others are in lower-case lettering. What is the difference between the two?**

Sequence in lower case has been identified by RepeatMasker as low-complexity or repetitive elements. You can find this description by clicking on the "Legend" link located above the sequence, which will take you to a sequence descriptions page. **(9/23/05)**

**I have come across nucleotide representations that I don't understand. What does "R" mean?**

"R" is part of the IUPAC code for nucleotide variations which represents "A" or "G". You can find all of the IUPAC nucleotide codes online. **(3/9/05)**

**I think "alleles" and "db SNP allele" may be switched in rs28944222, where dbSNP shows A/G; S, P; and in rs28944221 where dbSNP shows T/C; N, and D.**

Both of these rs numbers mapped to the reverse strand of the contig, while the mRNA mapped to the forward strand:

=======================> Contig [Forward]

-------------> mRNA [Forward]

<------ SNP [Reverse]

You must therefore use the complementing nucleotides of the SNP alleles in order to get the correct codon, which will in turn, code for the correct amino acid:

T/C is the complement of A/G and codes for S, P

A/G is the complement of T/C and codes for N, D (**1/5/06**)

**What method do you use to mark the position of an in/del (insertion/deletion)?**

We use "N' to indicate that a SNP is not a true single base nucleotide substitution. For example, "N" is used for an in/del SNP of any length, whether it is a microsatellite, or a multiple base substitution. (**2/28/05**)