|  |  |  |
| --- | --- | --- |
| **­Example** | **Key Format Features** | **Element** |
| *Specie­s name* or *S. name* | Genus (*“Species*”) is capitalizedspecies (“*name*”) is lowercase. italicized or underlined | Scientific name for an organism. A strain name may follow e.g. *E. coli* K12 |
| *genE* | lowercase for 3 letters, then uppercase. italicized or underlined | Genes (DNA).  |
| GenE | 1st letter is uppercase. Not italicized | Protein GenE is encoded by *genE* on DNA |
| *P*­*genE* | Capitol “P” = PromoterSubscript/italicized gene name | Promoter for *genE* |
| *P*­*genE*-*proT* |  | The promoter for *genE* is used to express the gene *proT*. This is a “transcriptional fusion” |
| pGenE | lowercase “p”  | A plasmid that has something to do with GenE (hopefully the paper will define it better) |
| Wildtype or WT |  | The “parent” version of the bacteria. Wildtype is somewhat randomly decided. But mutants are usually made by changing a WT in some way.  |
| Δ*genE* strain | Δ (capitol delta) | This is a mutant with *genE* deleted |
| *genE* strain | Gene name followed by “strain” | This is a strain with *genE* mutated in some way. This isn’t the most accurate notation but some people choose it because it helps readability.  |
| *genE*::Tn7 | Gene name::Tn7 | *genE*:: indicates that the *genE* gene sequence is interrupted by some other type of DNA sequence.“Tn” means transposonThis can refer to the actual mutation (*genE*::Tn7) or the strain (*genE*::Tn7 strain) |
| *genE+* | Gene with superscript + | Wildtype allele of the gene |
| *genE-* | Gene with superscript - | Mutant allele of the gene |
| KanR | Looks like a protein, but there’s something superscripted… | This is a phenotype. In this case, resistance to the antibiotic kanamycin.  |
| His+ | ^ | Also a phenotype. Phenotypes that involved the use of a nutrient source (e.g. Histidine) are usually given in +/- notation |

Less common notations that you will probably run into:

|  |  |
| --- | --- |
| ‘GenE | The apostrophe indicates that the protein is missing the N-terminus (beginning) |
| GenE’ | Here the apostrophe indicates that the protein is missing the C-terminus (end) |
| ΔCT or ΔNT | This is very short notation that is sometimes used to indicate that the DNA encoding C-terminus or N-terminus of a protein are deleted |
| D572E | This indicates a missense mutation where the DNA that originally encodes a D (aspartate)\* at position 572 now encodes an E (glutamate)\*. This could be in the context of natural variation of alleles or a direct mutants the authors create to test a hypothesis. \*Learn your 1 letter amino acid codes! They’re used a lot in molecular biology papers. ☺ |

More examples on Wikipedia!

<http://en.wikipedia.org/wiki/Gene_nomenclature#Bacterial_genetic_nomenclature>

### Mutant nomenclature[[edit](http://en.wikipedia.org/w/index.php?title=Gene_nomenclature&action=edit&section=9)]

If the gene in question is the wildtype a superscript '+' sign is used:

* *leuA+*

If a gene is mutant, it is signified by a superscript '-':

* *leuA-*

By convention, if neither is used, it is considered to be mutant.

There are additional superscripts and subscripts which provide more information about the mutation:

* ts = temperature sensitive (*leuAts*)
* cs = cold sensitive (*leuAcs*)
* am = amber mutation (*leuAam*)
* um = umber (opal) mutation (*leuAum*)
* oc = ochre mutation (*leuAoc*)
* R = resistant (*rifR*)

Other modifiers:

* Δ = deletion (Δ*leuA*)
* - = fusion (*leuA*-*lacZ*)
* : = fusion (*leuA*:*lacZ*)
* :: = insertion (*leuA*::Tn*10*)
* Ω = a genetic construct introduced by a two-point crossover (Ω*leuA*)
* Δ*gene name*::*drug resistance marker* = deletion/replacement (Δ*leuA*::*npt* indicates that the *leuA* gene has been deleted and replaced with the gene for neomycin phosphotransferase, which confers resistance to kanamycin)