Editing Guidelines		
Status: Published		Document ID: 1163
DATE EFFECTIVE 07/08/2021	APPROVED BY mtDNA Technial Leader	PAGE 1 OF 2

1 Editing Guidelines

1.1 Sequencher base calls can be modified if the underlying data support it. The analyst can change an "N" call into a base determination, change a called base to an "N" call, insert an additional base, remove a position, or trim a sequence. A base call must not be edited without proper justification.

1.2 The following editing codes are used to specify reasons for each type of edit:

1.2.1 Editing code 1. Base is inserted by the software due to broad peak, peak artifact, or analysis default spacing; this extra base is removed by the analyst.

1.2.2 Editing code 2. (i) Base is omitted by the software; base is inserted and called by the analyst because an authentic peak is present or it is needed to correct for spacing, or (ii) Base is called "N" by the software; base is called by the analyst because an authentic peak is present.

Discussion: This editing code will usually be documented as a 2,3 since neighboring peak artifacts (e.g. peak spacing) are usually present in these instances (see also discussion in editing code 3). The only time a 2 code will be used on its own is when a peak in question has absolutely no overlap with either of its neighboring peaks; these instances are rarely seen. In instances when a base is being called to correct for spacing, however, and an authentic peak is not present (e.g. a broad peak due to the presence of multiple identical bases), the call being made by the analyst will be an "N" call.

1.2.3 Editing code 3. Dye artifact, electrophoretic artifact, or neighboring/artifactual peak interference is present. Base is called incorrectly by the software due to the artifacts mentioned; base call is corrected by the analyst.

Discussion: This editing code may be combined with the 2 code if it describes why the software would have omitted a base or called it an "N" (e.g. 2,3 covers what you see, why you see it, and what you did). For all other instances where the analyst is correcting a call, the 3 code will be used on its own. An example of this is in instances of length heteroplasmy when the software is incorrectly calling the smaller (e.g. artifactual) peak of the two peaks present at a given position.

1.2.4 Editing code 4. Trimmed to remove end sequence.

1.2.5 Editing code 5. Changed to IUPAC code due to sequence heteroplasmy.

1.2.6 Editing code 6. Changed to "N" due to ambiguous base.

Discussion: This code is an all-encompassing one to be used whenever the analyst is changing a call made by the software to an "N". Examples when to use this code include

instances of length heteroplasmy when the interfering artifactual peak is either overlapping or higher in amplitude than the authentic peak. This code can also be used for those instances where there may be three or more peaks at a given position resulting either from multiple length heteroplasmy or a sequence artifact that results in multiple peaks at one position.

1.2.7 Editing code 7. Other.

Discussion: Editing for any other reasons than described in editing codes 1-6 should be documented with a comment explaining the edit.

1.3 Many software calls can be easily resolved and corrected by the analyst. However, ambiguous situations should not be edited. If an electrophoresis problem is suspected, this sample should be re-injected. Sequence information at each base position should be confirmed by data from both DNA strands when possible. Single-stranded regions present due to length heteroplasmy, must be confirmed by confirmatory sequencing of the same strand in the same direction. The Sequencher complementary strand alignment will flag conflicts between the two sequencing directions for all strands imported into the contig.