# ori: there is not exactly 1 occurence of the origin sequence

**Reference**

full genome sequence

origin sequence

genome

1    L

L  1

**No errors**

genome

1    L

origin sequence

L  1

OR

L  1

origin sequence

**ori(origin)**
sequence does not
contain exactly
1 occurence of the
origin sequence

genome

1    L

L  1

OR

origin sequence

L  1

origin sequence
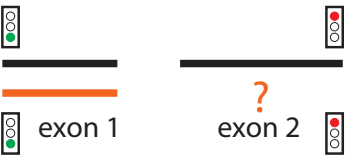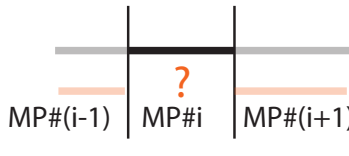
Example output from annotation script:

```
HQ693428    N/A  N/A          ori  there is not exactly 1 occurrence of origin sequence [0 occurrences]
```
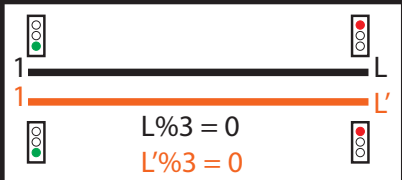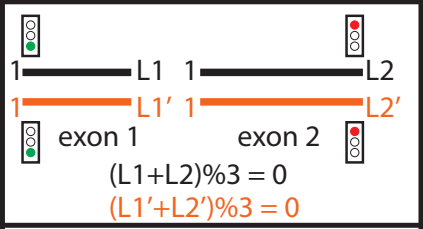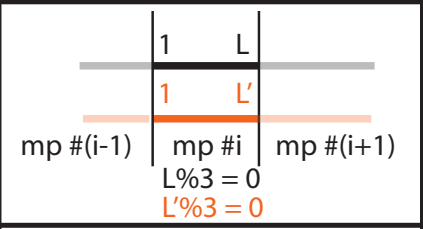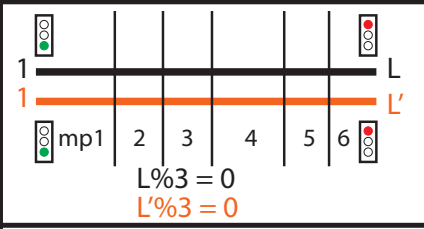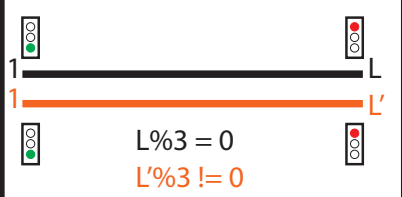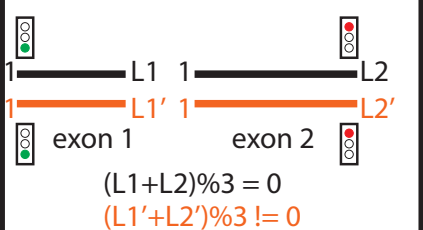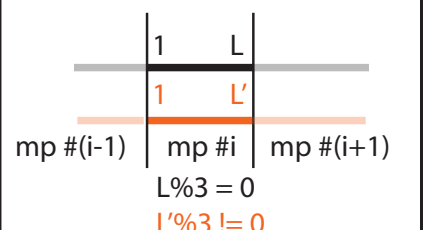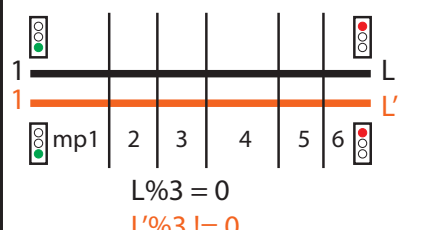
# nop: unable to identify homologous feature



Example output from annotation script:

```
AM421536     1  CDS#1   nop  unable to identify homologous feature
```
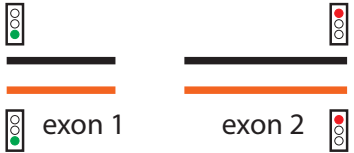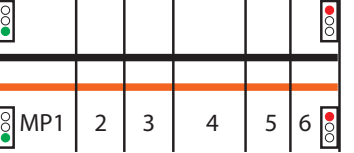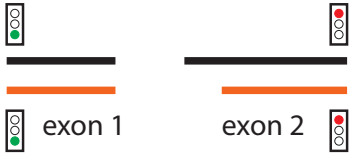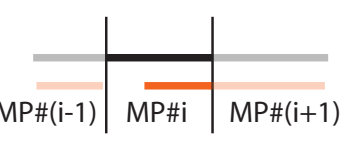
# nm3: length of nucleotide feature is not a multiple of 3

| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors | $L\%3 = 0$ <br> $L'\%3 = 0$ | exon 1    exon 2 <br> $(L1+L2)\%3 = 0$ <br> $(L1'+L2')\%3 = 0$ | mp #(i-1)   mp #i   mp #(i+1) <br> $L\%3 = 0$ <br> $L'\%3 = 0$ | mp1 2 3 4 5 6 <br> $L\%3 = 0$ <br> $L'\%3 = 0$ |
| nm3 (not a multiple of 3) <br> length is not a multiple of 3 | $L\%3 = 0$ <br> $L'\%3\ != 0$ | exon 1    exon 2 <br> $(L1+L2)\%3 = 0$ <br> $(L1'+L2')\%3\ != 0$ | mp #(i-1)   mp #i   mp #(i+1) <br> $L\%3 = 0$ <br> $L'\%3\ != 0$ | mp1 2 3 4 5 6 <br> $L\%3 = 0$ <br> $L'\%3\ != 0$ |

Example output from annotation script:

```
JF920397      2  MP#2    nm3  length of nucleotide feature is not a multiple of 3 [299]
```
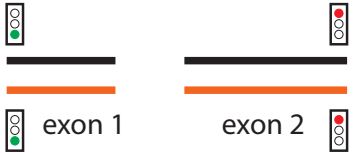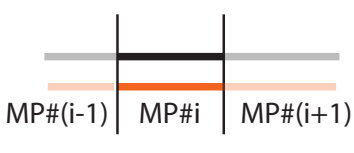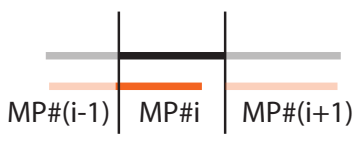
# bd5: alignment to reference does not extend to 5' end of reference

| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| **No errors** | | exon 1  exon 2 | MP#(i-1)  MP#i  MP#(i+1) | MP1  2  3  4  5  6 |
| **bd5 (5' boundary)** alignment does not extend to 5' boundary of reference | ? | exon 1  exon 2 | MP#(i-1)  MP#i  MP#(i+1) | N/A |

Example output from annotation script:

```
AY858050      2  MP#2    bd5  alignment to reference does not extend to 5'
                              boundary of reference [10 nt from 5' end]
```

# bd3: alignment to reference does not extend to 3′ end of reference

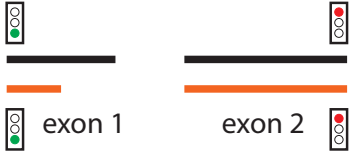| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors | | | MP#(i-1) \| MP#i \| MP#(i+1) | MP1 \| 2 \| 3 \| 4 \| 5 \| 6 |
| | | exon 1 \| exon 2 | | |
| **bd3 (3′ boundary)** alignment does not extend to 3′ boundary of reference | ? | | MP#(i-1) \| MP#i \| MP#(i+1) | N/A |
| | | exon 1 \| exon 2 | | |

Example output from annotation script:

```
KM198581      9   CDS#3      bd3  alignment to reference does not extend to 3'
                                  boundary of reference [66 nt from 3' end]
```

# olp: lack of an expected overlap with another feature
# or additional unexpected overlap with another feature



Example output from annotation script:

```
FJ562227    6  CDS#6      olp  feature does not overlap with same set of features as in reference [-(6.1,1.1),-(6.1,7.1)]
```

# str: predicted start of CDS is not valid start codon

| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors | | | MP#(i-1)   MP#i   MP#(i+1) | MP1  2  3  4  5  6 |
| str (start)<br>predicted CDS start<br>is not a valid start codon | | exon 1    exon 2 | N/A | MP1  2  3  4  5  6 |

Example output from annotation script:

```
KR006962      3  CDS#3      str  predicted CDS start position is not beginning of ATG start codon
                                  [TNN starting at position 4472 on strand +]
```
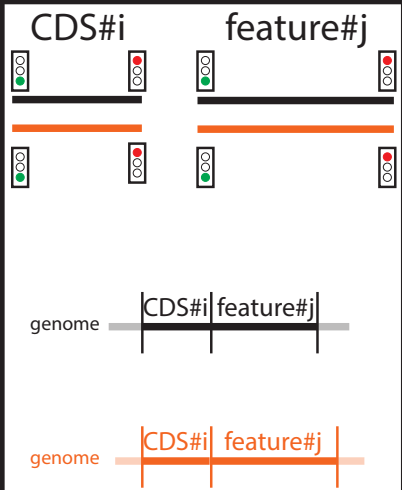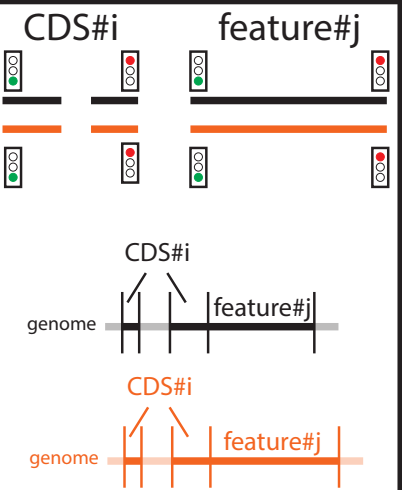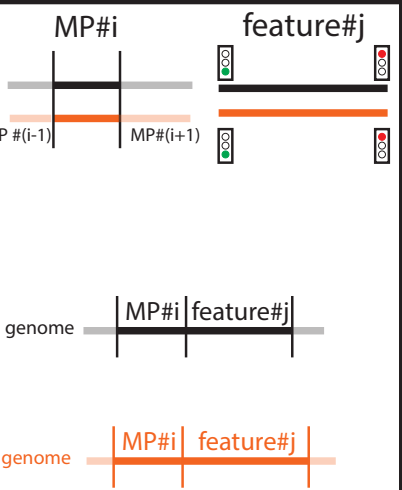
# stp: predicted stop of CDS is not valid stop codon
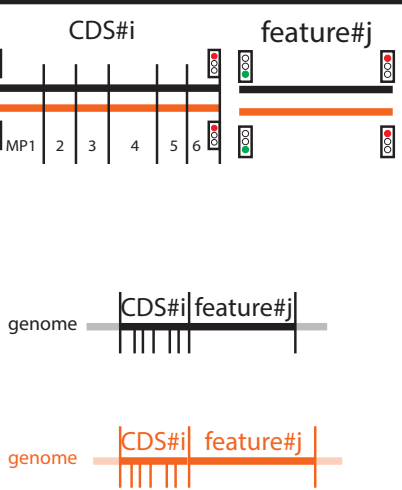
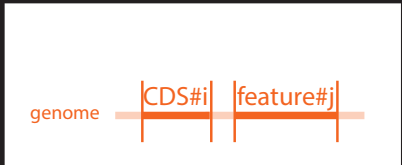| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors | | | | |
| stp (stop) predicted CDS stop is not a valid stop codon | | | N/A | |

Example output from annotation script:

```
GQ377530      6  CDS#6   stp  predicted CDS stop  position is not end of valid stop codon
                               (TAG|TAA|TGA) [CAG ending at position 2452 on + strand]
```

# ajb: MP or CDS is not adjacent to same set of features before it as in the reference



Example output from annotation script:

```
KC131141    11  MP#11   ajb feature (MP or CDS) is not adjacent to same set of
                            features before it as in reference [-(11.1,10.1)]
```

# aja: MP or CDS is not adjacent to same set of features after it as in the reference
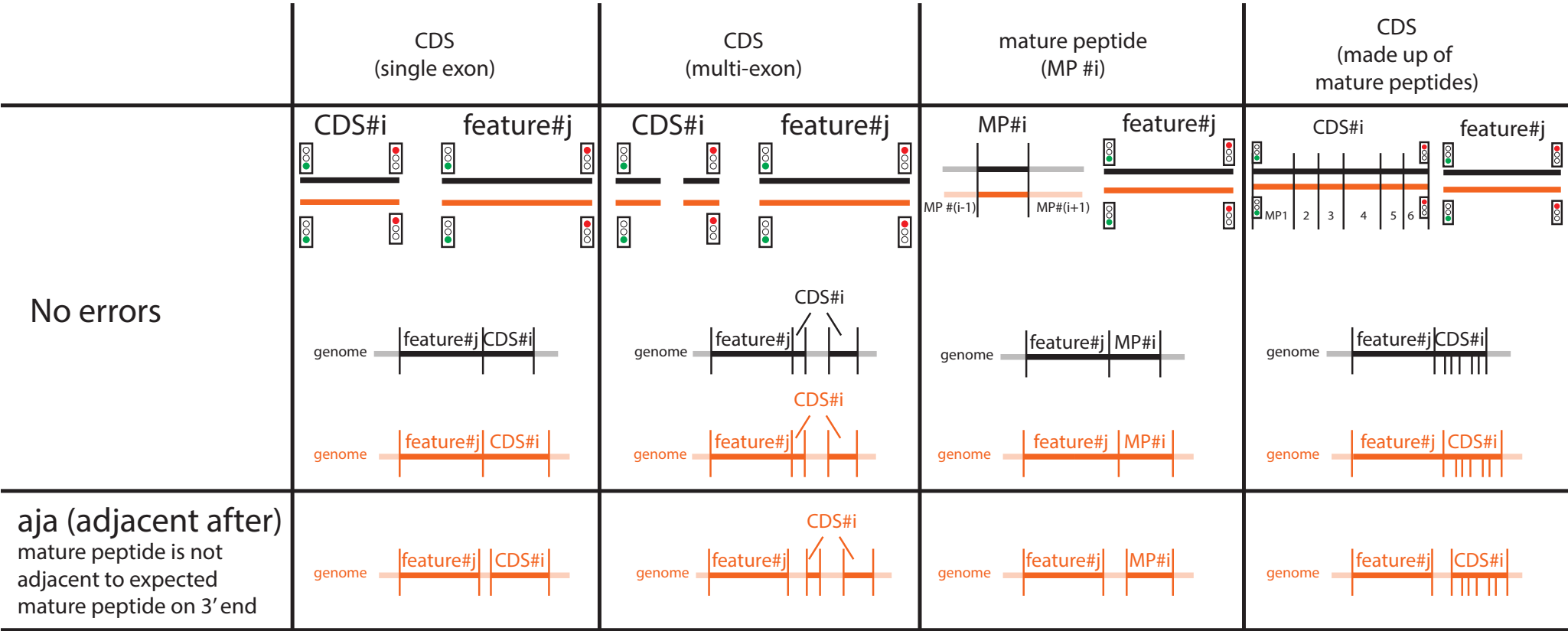


Example output from annotation script:

```
CS477302     10  MP#10  aja  feature (MP or CDS) is not adjacent to same set of
                                 features after it as in reference [-(10.1,11.1)]
```

# trc: truncation due to an early stop



Example output from annotation script:

```
EU628610      3  CDS#3       trc  in-frame stop codon exists 5' of stop position predicted by homology to reference
                                  [homology search predicted 1796..1356 exon 2 of 2 revised to 1796..1371 (stop shifted 15 nt)]
```
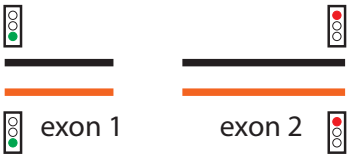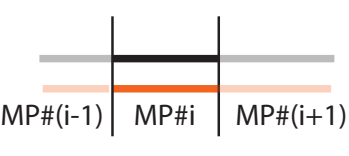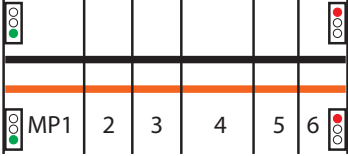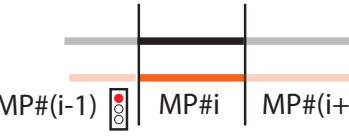
# ext: extended feature due to a missing stop codon



Example output from annotation script:

```
HM631854     14  MP#14       ext  first in-frame stop codon exists 3' of stop position predicted by homology to reference
                                   [homology search predicted 7544..10230 revised to 7544..10288 (stop shifted 58 nt)]

HM631854     15  CDS(MP)#1   ext  first in-frame stop codon exists 3' of stop position predicted by homology to reference
                                   [homology search predicted 74..10233 revised to 74..10291 (stop shifted 58 nt)]
```

# ntr: mature peptide is not translated because of upstream stop

| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors |  |  exon 1 / exon 2 |  MP#(i-1) / MP#i / MP#(i+1) |  MP1 2 3 4 5 6 |
| ntr (not translated) mature peptide is not translated due to early stop 5' of predicted start | N/A | N/A |  MP#(i-1) / MP#i / MP#(i+1) | N/A |

Example output from annotation script:

```
KC131141     11  MP#11   ntr  mature peptide is not translated because its CDS has an in-frame stop 5' of the mature
                              peptide's predicted start [early stop in mature peptide 10 ending at position 6274]
```
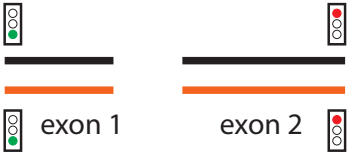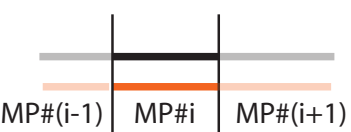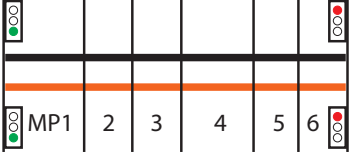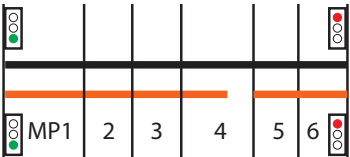
# nst: no in-frame stop codon exists 3' of predicted valid start codon



Example output from annotation script:

```
HQ234499      15  CDS(MP)#1   nst   no in-frame stop codon exists 3' of predicted valid start codon [inferred stop
                                    codon position (3 nt 3' of 10269 on + strand) is off the end of the sequence]
```

# aji: CDS has at least one inconsistent MP adjacency relative to the reference
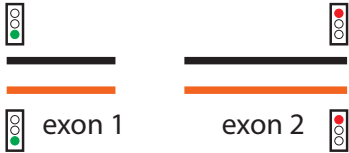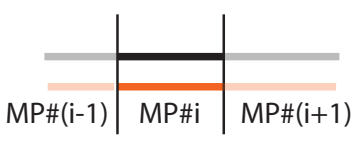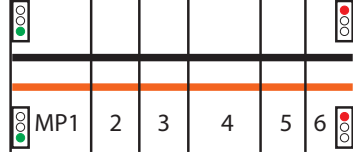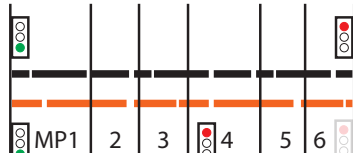
| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| **No errors** |  |  | MP#(i-1) MP#i MP#(i+1) | MP1 2 3 4 5 6 |
| **aji (adjacency inconsistency)** CDS has at least one inconsistent adjacency relative to reference | N/A | N/A | N/A | MP1 2 3 4 5 6 |

Example output from annotation script:

```
HQ166035      15  CDS#1   aji   CDS comprised of mat_peptides has at least one adjacency inconsistency between
                            2 mat_peptides [MP#9 (4074..4463) not adjacent to MP#10 (4465..4572)]
```

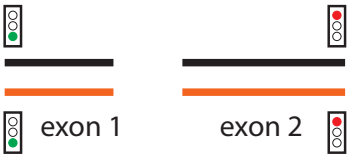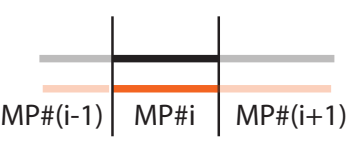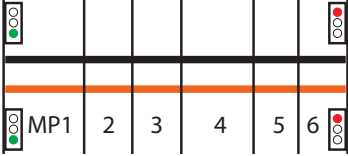# int: CDS has at least one MP that is not translated due to an early stop
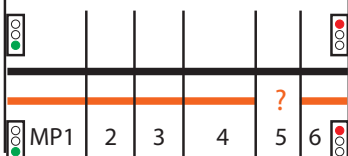
| | CDS (single exon) | CDS (multi-exon) | mature peptide (MP#i) | CDS (made up of mature peptides) |
|---|---|---|---|---|
| No errors |  |  exon 1 exon 2 |  MP#(i-1) MP#i MP#(i+1) |  MP1 2 3 4 5 6 |
| **int (interrupted translation)** CDS has at least one mature peptide that is not translated due to early stop | N/A | N/A | N/A |  MP1 2 3 4 5 6 ? |

Example output from annotation script:

```
JN975491      7  CDS(MP)#1   int   CDS comprised of mat_peptides is incomplete: at least one primary mat_peptide
                                    is not translated due to early stop (ntr) [MP#2, MP#3, MP#4, MP#5, MP#6]
```

# inp: CDS has at least one MP that is not identified



Example output from annotation script:

```
KC407671     16  CDS(MP)#1   inp  CDS comprised of mat_peptides is incomplete: at least one
                                   primary mat_peptide is not identified (nop) [MP#1, MP#3]
```