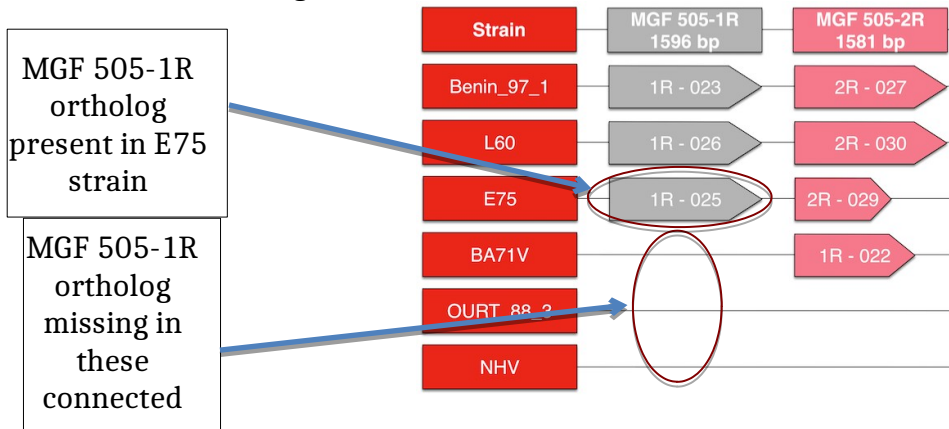


# MGF 505 Figure Legend



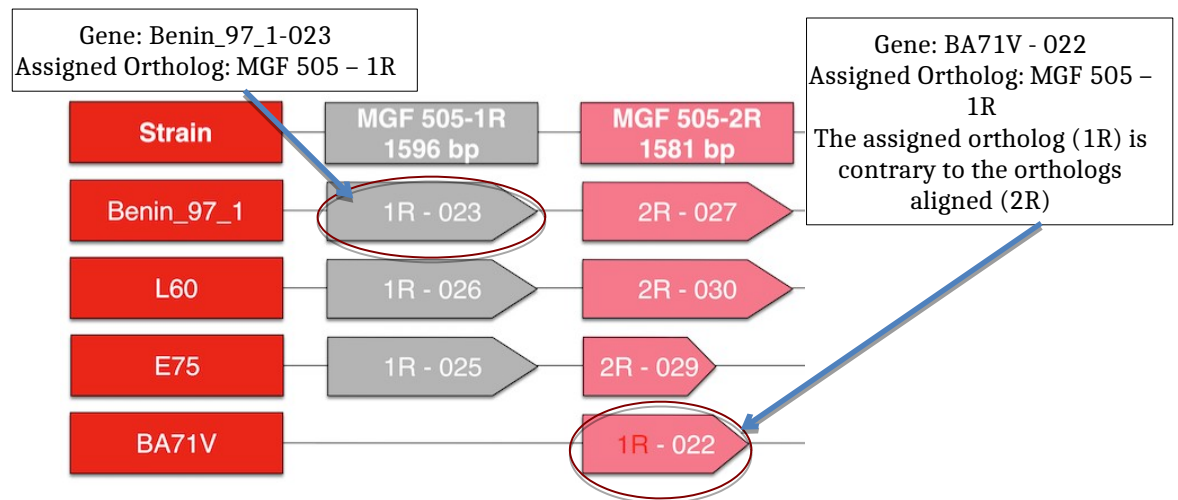
## Ortholog columns and gene size:

All genes of same width as the heading are approximately the length of the reference length stated.



## Genes:

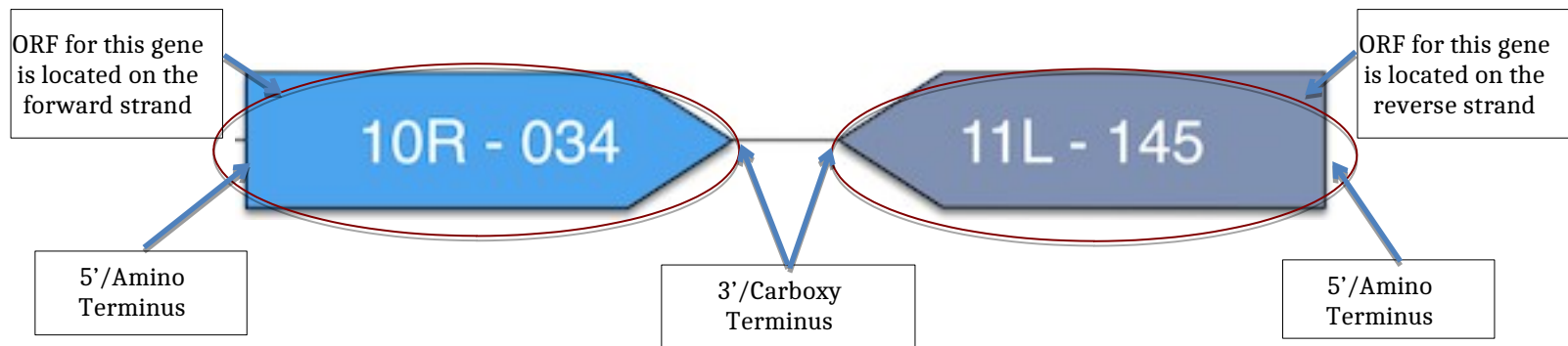
The above figure indicates MGF 505-1R ortholog is missing in strains BA71V, OURT\_88\_3, and NHV but present in Bening\_97\_1, L60, and E75



## Gene labels:

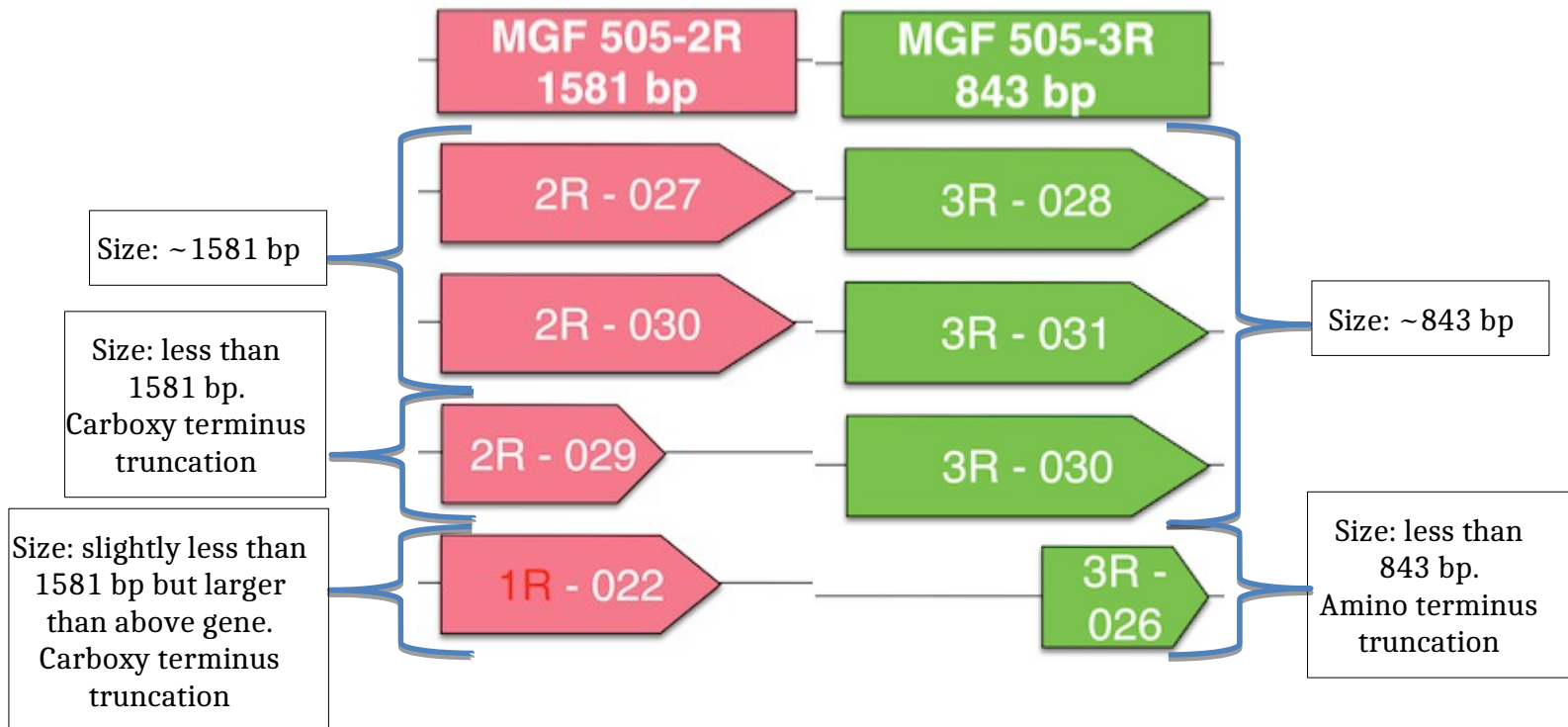
The annotation of each gene is in two parts: first, the ortholog group originally assigned during initial annotation, followed by the corresponding gene number of the connected strain. If the currently assigned ortholog group is labelled in **red**, this indicates that the currently assigned ortholog is contrary to orthologs it aligns to within this diagram.

# MGF 505 Figure Legend



## Gene Orientation:

MGF 300 has “R” orthologs that are transcribed on the forward strand (5' → 3') and “L” orthologs that are transcribed on the reverse strand (3' ← 5'). “R” gene boxes are pointed to the right and “L” gene boxes are pointed to the left.



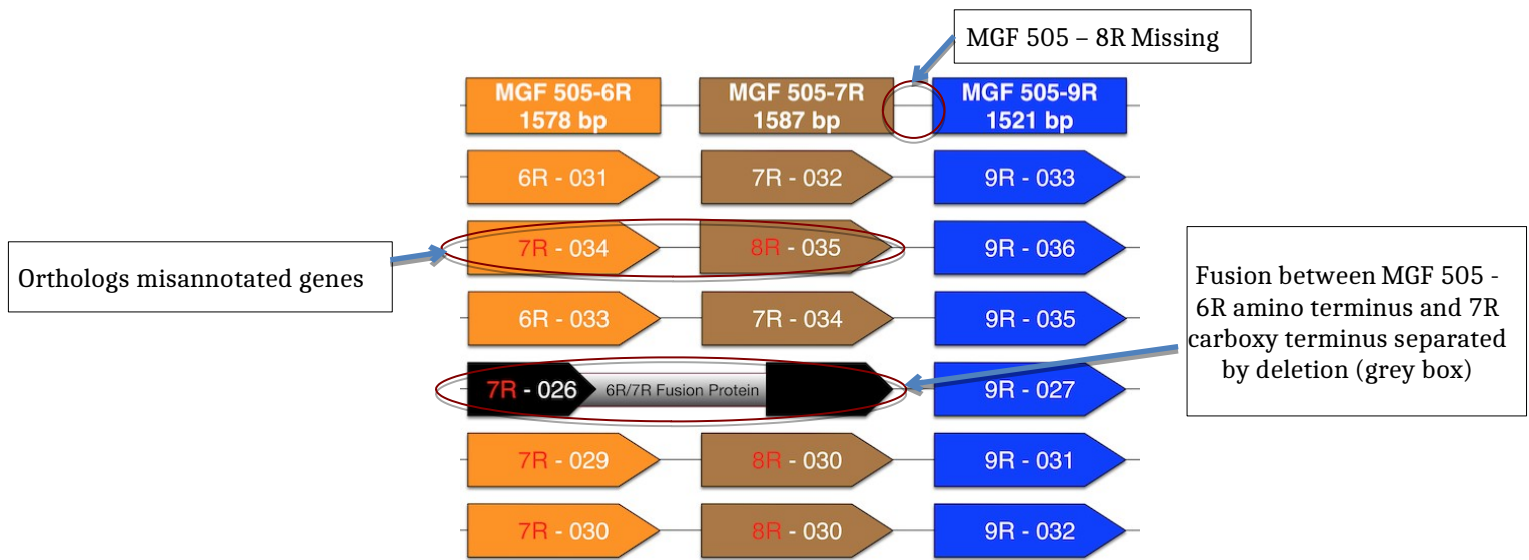
## Relative gene size and truncations:

A smaller or larger gene box indicates the size difference of the gene relative to other genes of the same ortholog

A 5'/amino terminus that is not aligned with the amino terminus of the full-length genes indicates an amino terminus truncation of this gene.

A 3'/carboxy terminus that is not aligned with the carboxy terminus of the full-length genes indicates a carboxy terminus truncation of this gene.

# MGF 505 Figure Legend



## MGF 505 Fusion:

Genes encoded by an open reading frame that aligns across multiple ortholog loci due to genomic deletions are labelled in the above diagram as “fusion” genes and are represented by black gene boxes. The size and alignment of the black gene boxes to a given ortholog group is representative of how much of the ortholog is fused and which region.

The fused ortholog groups are labelled along the deletion box connecting the fragments.

## MGF 505-8R Missing:

MGF 505-8R is only annotated in L60, OURT\_88\_3, and NHV. There is high similarity between the MGF 505-6R and 7R orthologs and the ortholog assignment of the fusions between these orthologs as MGF 505-7R is most likely the cause of the 8R group. If MGF 505-8R is present in a connected genome, MGF 505-6R is skipped. Due to only 3 genomes containing the 8R ortholog compared to 8 genomes containing the 6R ortholog, the use of the 6R ortholog is favoured in assigning orthologs in this diagram.