

Following a gene duplication, what is the most common outcome?

- A) Neofunctionalization
- B) Decay of one of the genes (non-functionalization)
- C) Subfunctionalization

What is the most common pathway for gene family expansion in prokaryotes:

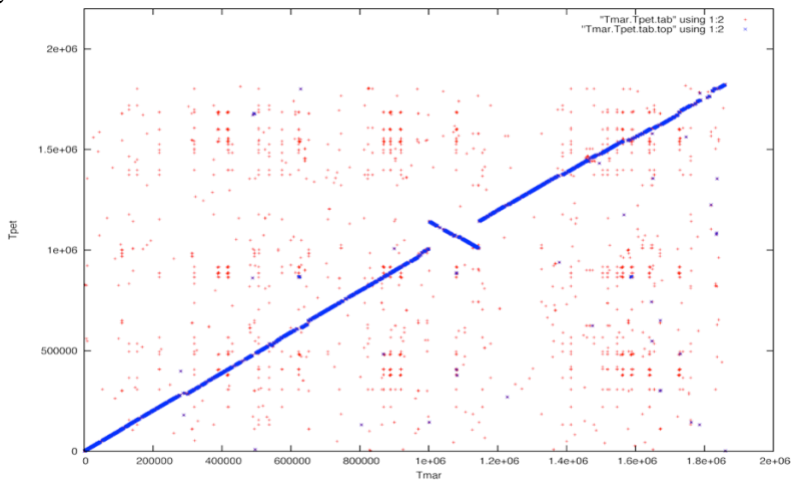
- A) Gene duplications
- B) Genome duplications
- C) Gene transfer
- D) De-novo creation of genes from non-coding sequences?

Why is percent identity in blast search results not a good measure to assess significant sequence similarity?

Two BLASTp hits have E-values of 0.0 and $1e-67$. Which of these statements is true?

- A) $1e-67$ is a very large negative number, giving a highly significant match
- B) E-values of 0 aren't possible
- C) $1e-67$ is a very small positive number, indicating a significant match
- D) an unexpected intein probably corrupted the PSSM, and it zeroed out
- E) these proteins almost certainly are ATPase subunits from an extinct Archaeon
- F) you forgot to format the database first (makeblastdb), a common mistake
- G) the only way to tell if these hits are significant is to look at the percent identity of the high-scoring segment pairs (HSPs)
- H) $1e-67$ looks like an isoelectric point, meaning that the query sequence is likely from an extremophilic organism living in deep-sea hydrothermal vents
- I) these values should be interpreted as 0% and 67% homology, respectively, indicating that horizontal transfer occurred recently

In the following gene plot. The axes give the genes locations in the *Thermotoga maritima* and *Thermotoga petrophila* genomes, respectively. Red: all significant blast hits, blue top scoring blast hits only.



What mechanism is this graph depicting when blue dots appear on the downward sloping diagonal?

- A. Recombination between two points in the genome leading to an inversion
- B. Deletion or Insertion
- C. Translocation
- D. Paralogs
- E. All of the above

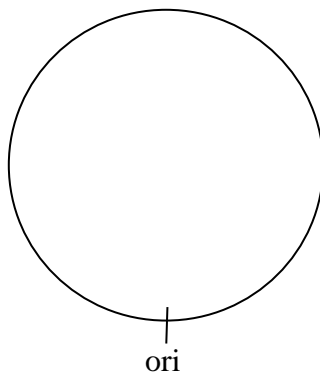
What does the blue line represent?

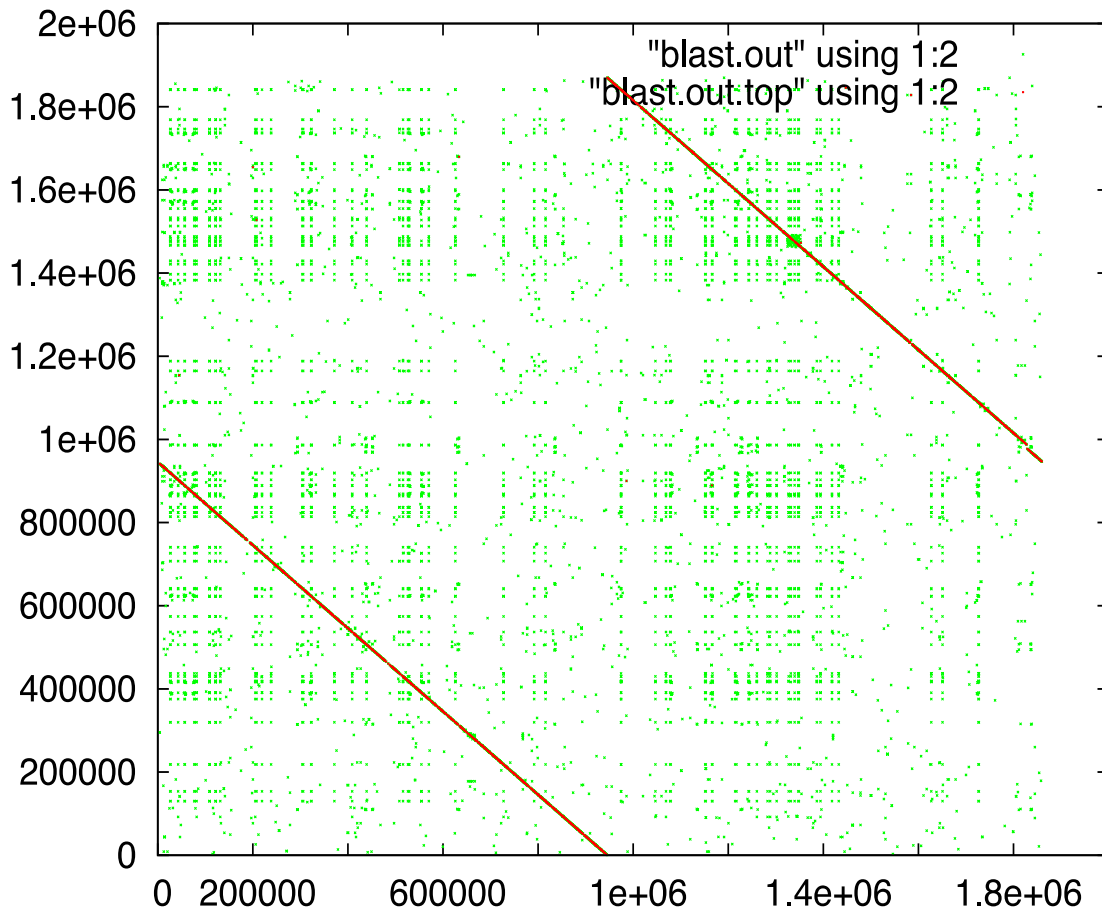
- A. The location of all genes in one genome versus the location of the top scoring blast hit in the other genome
- B. The location of all genes in one genome versus the location of ALL the blast hits in the other genome
- C. The location of the gene in the environment
- D. A and B
- E. None of the above

What do the red dots represent?

- A. False positives
- B. False negatives
- C. Partial matches
- D. Paralogs in the genome
- E. Orthologs in the genome

Map the genome rearrangement shown onto the circular genome below:

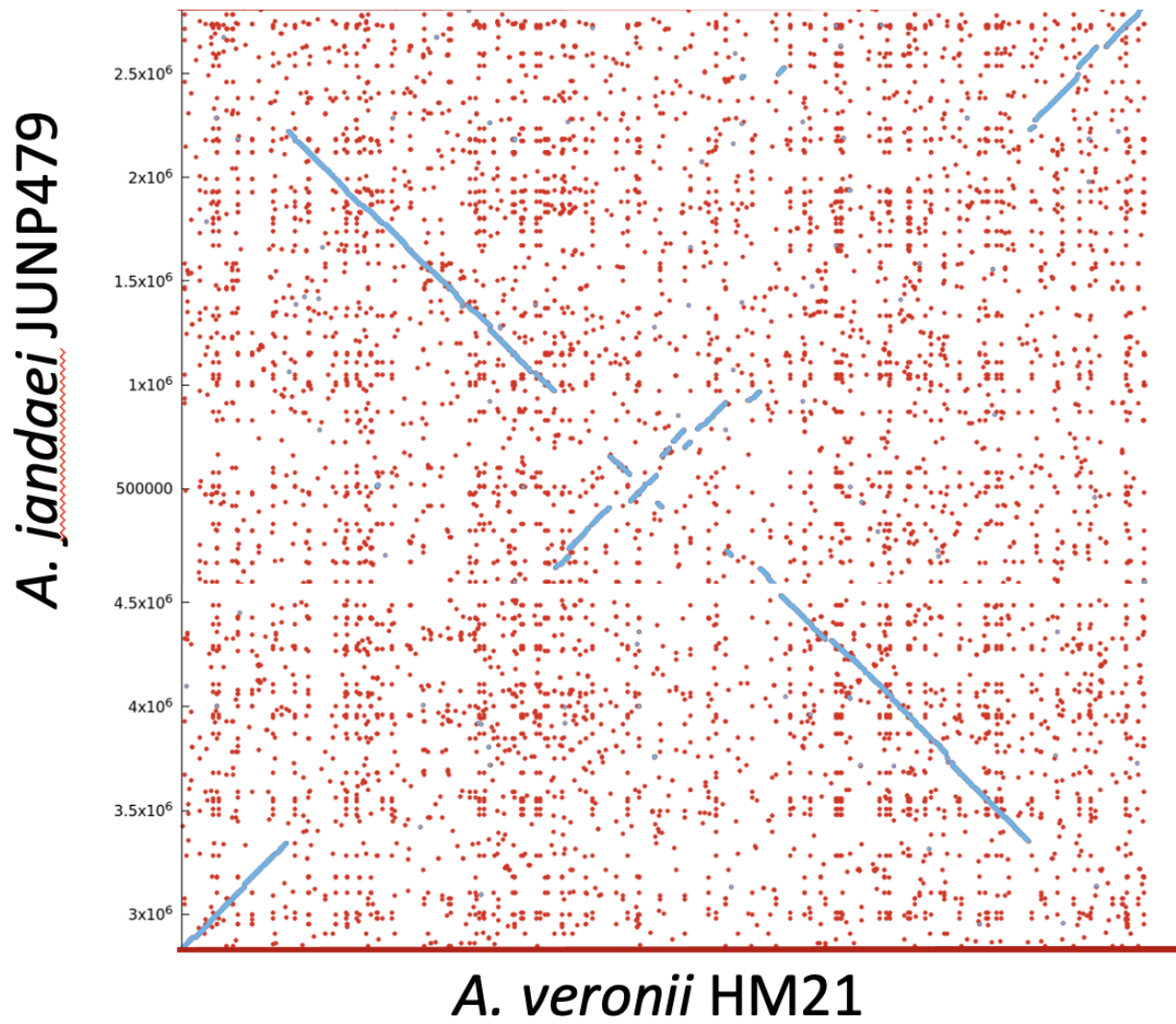




What might have happened in the genome plot given above?

- A. One massive genome inversion, involving half of the genome
- B. One round of whole genome duplication, so that every gene is present in one of the two genomes twice and only once in the other
- C. Two identical genomes were used, but the origin of replication was miscalled in one.
- D. There is a strong strand bias.
- E. The “forward” strand of one genome was reported, while the reverse complement was reported in the other genome

Which recombination events could have given rise to the following gene plot (ignore the many tiny events close the terminus of replication):



Map the genome rearrangement shown onto the circular genome below:

